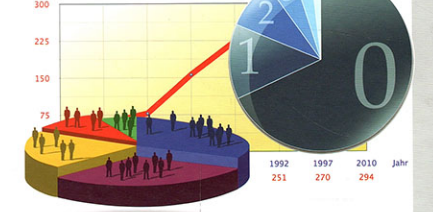
طرق الإقتصاد القياسي

محاضرات وتطبيقات



الدكتور شيخي محمد





طرق الاقتصاد القياسي

محاضرات وتطبيقات

طرق الاقتصاد القياسي

محاضرات وتطبيقات

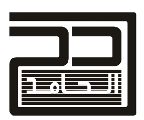
الدكتور

محمد شيخي

أستاذ باحث بجامعة ورقلة، الجزائر

الطبعة الأولى

2011م



المحتويات

الموضـوع	الصفحة

الموضـوع	الصفحة

الموضـ . . وع	الصفحة

الموضـــوع	الصفحة

الموضـوع	الصفحة

الموضـوع	الصفحة

الموضـوع	الصفحة

تقديم

يعد الاقتصاد القياسي أحد فروع العلوم الاقتصادية المستخدمة للأساليب الكمية في تحليل الظواهر الاقتصادية، فلقد ساعد التطور في النظرية الإحصائية والاقتصادية وثورة المعلومات على حدوث تطور كبير في مجال الاقتصاد القياسي خلال فترة زمنية قصيرة.

لقد استُخدم لفظ اقتصاد قِياسي Econometrics لأول مرة سنة 1926. إن أصل هذا المصطلح Econometrics يوناني و يتكون من مقطعين هما Economic أي علم الاقتصاد و Metrics أي القياس (المتر) ويعرفه البعض بأنه القياس في الاقتصاد، وبصورة أكثر تفصيلا هو العلم الذي يهتم بقياس العلاقات الاقتصادية من خلال بيانات واقعية، بغرض اختبار مدى صحة هذه العلاقات كما تقدمها النظرية، أو تفسير بعض الظواهر، أو رسم بعض السياسات، أو التنبؤ بسلوك بعض المتغيرات الاقتصادية.

إن غالبية العلاقات التي تقدمها لنا النظرية الاقتصادية، يمكن صياغتها في صورة نماذج رياضية تُقدَّر من واقع البيانات الفعلية، وهذا يُمكّننا من وضع تنبؤات على الآثار الكمية على أحد المتغيرات الاقتصادية التي يمكن أن تترتب على التغير في أحد أو بعض المتغيرات الاقتصادية الأخرى. وحيث أن أغلب المتغيرات الاقتصادية قابلة للقياس الكمي مثل السعر، الدخل..الخ، فإنه يمكن استخدام الأسلوب الرياضي في شرح العلاقات الاتجاهية، كما تحددها النظرية الاقتصادية، بين هذه المتغيرات. كما أن هناك نماذج أخرى غير سببية تعتمد على القيم التاريخية للمتغير المراد التكهن بقيمته المستقبلية ولا تحتاج إلى تحديد المتغيرات التي تفسر سلوكه ويعرف هذا النوع من النماذج باسم نماذج السلاسل الزمنية التي تذبذباتها ناتجة عن الاتجاه العام الذي يعبر عن الحركة طويلة المدى والتقلبات الموسمية والدورية وكذلك التقلبات العشوائية.

هناك علاقة بين الاقتصاد القياسي والفروع الأخرى، حيث أن الإحصاء يمدنا بأساليب وطرق القياس مثل الارتباط والانحدار، بالإضافة إلى البيانات الواقعية المبوبة؛ أما النظرية الاقتصادية فتُحدد لنا العلاقات الاقتصادية المراد قياسها من خلال الفرضيات التي تقدمها، بينما يضع الاقتصاد الرياضي هذه العلاقات النظرية في صيغ رياضية هي المعادلات التي تأخذ أشكالا دالية مختلفة و قابلة للقياس. الهدف الأساسي إذن هو بناء النماذج القياسية الاقتصادية في شكل قابل للاختبار الميداني و تقدير واختبار هذه النماذج مستعملين البيانات المتوفرة ثم في الأخير استعمال النماذج المقدرة لغرض التنبؤ، التحليل الاقتصادي أو اتخاذ القرارات المناسبة.

تقترح الطبعة الأولى من هذا الكتاب دروسا في الاقتصاد القياسي مع إعطاء أمثلة محلولة بطريقة عملية و بيداغوجية لتمكين الطالب من التحكم الجيد في القوانين المعطاة في المحاضرة من جهة ومن جهة أخرى تشجيع الطالب على استخدام برمجيات الاقتصاد القياسي التي تعتبر كضرورة ملحة في وقتنا الحاضر. سيتم حل بعض الأمثلة باستعمال برمجيات RATS 5.04 و Eviews 4.0 و GAUSS وبرمجيات أخرى خاصة.

هذا الكتاب موجه لطلبة السنة الثالثة علوم اقتصادية و تجارية وعلوم التسيير بنظاميها الكلاسيكي وال .LMD وطلبة السنة الأولى ماستر كما يخص أيضا طلبة الدكتوراه والباحثين في الميدان. يتميز هذا الكتاب بسهولة معالجة الموضوعات واستخدام الأساليب الرياضية الأولية، ثم هناك بعض البراهين في كل فصل لإثبات النظريات والعلاقات الرياضية المعقدة التي وردت به (يمكن للطالب العادي إهمالها).

تحتوي الطبعة الأولى من هذا الكتاب على تسعة فصول، سنتطرق في الفصلين الأول والثاني إلى طرق تحليل الانحدار الخطي البسيط والمتعدد و الفصل الثالث يتناول مشاكل القياس الاقتصادي التي تتعلق باختلال في فرضيات النموذج. يحتوي الفصل الرابع على طرق أخرى في تحليل الانحدار مثل الانحدار غير الخطي وطريقة النواة، أما الفصل الخامس

فنعطي فيه نبذة عن نماذج المعادلات الآنية وطرق تقديرها. تأتي بعد ذلك الفصول المتبقية لتعالج بتعمق السلاسل الزمنية العشوائية وهو المجال الذي زاد استخدامه كثيرا في الآونة الأخيرة (دراسة الخصائص الإحصائية للسلسلة، نماذج ARMA، نماذج VAR ومشكل التكامل المشترك، نماذج ARCH، الشواش والذاكرة الطويلة والسيرورات غير المعلمية).

الفَصْلُ الأَوَّلُ

تحليل الانحدار الخطي البسيط

الفصل الأول

تحليل الانحدار الخطي البسيط

يعتبر الانحدار الخطي البسيط أبسط أنواع نماذج الانحدار، بحيث يوجد العديـد مـن العلاقات الاقتصادية التي يمكن قياسها باستخدام هذا الأسلوب، مثـل علاقـة الإنفـاق الاستهلاكي والدخل المتاح، وعلاقة الكمية المطلوبة من السلعة وسعرها، وأيضا مستوى البطالة مع معدل التضخم[1]... سنتطرق إذن في هذا الفصـل إلى تحليـل الانحـدار ذي متغيرين. نعطي أولا الصيغة الرياضية لهذا النموذج مع الفرضيات الأساسية حول الخطـأ العشوائي ثم في الفقرة الثانية من هذا الفصل نقوم بتعريف طريقة المربعات الصغرى العادية قصد تقدير معالم النموذج ودراسة خصائص المقدرات مع تشتتاتها وفي الجـزء الثـاني، سنتناول دراسة التوزيع الاحتمالي للمقدرات وبناء فترات الثقة قصد اختبار الفرضيات.

1. كتابة النموذج الخطي والفرضيات الأساسية:

يمكن نمذجة العلاقة بين المتغيرين Y_i و X_i على الشكل:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1.........n$$

حيث: Y_i يسمى بالمتغير المُفَسَّر أو التابع و X_i بالمتغير المُفَسِّر أو المستقل، β_0 و β_1 هما معلما النموذج.

أما ε_i فيمثل الخطأ في تفسير Y_i، ومنـه يمكـن كتابتـه انطلاقـا مـن العلاقـة:

$$\varepsilon_i = Y_i - \beta_0 - \beta_1 X_i$$

ويرجع وجود حد الخطأ إلى عدة أسباب منها:

- إهمال بعض المتغيرات المستقلة التي يمكن أن تؤثر على المتغير التابع في النموذج.
- الصياغة الرياضية غير السليمة للنموذج.
- حدوث خطأ في كل من تجميع البيانات وقياس المتغيرات الاقتصادية.

1- عبد القادر محمد عبد القادر، 1990، ص 89.

ويترتب على إسقاط هذا الافتراض حدوث أخطاء تحديد تتمثل فيما يلي:

- تحديد خاطئ للمتغيرات المستقلة: ويتمثل ذلك في إغفال متغيرات مستقلة هامة في نموذج الانحدار المراد تقديره، أو احتواء هذا النموذج على متغيرات مستقلة غير هامة.
- تغير معاملات الانحدار: إن معاملات الانحدار قد لا تظل ثابتة أثناء الفترة الزمنية التي تم تجميع البيانات عنها.
- العلاقة الحقيقية بين المتغير التابع والمستقل قد تكون غير خطية.

فرضيات النموذج:

أ. الفرضية الأولى: الأمل الرياضي للأخطاء معدوم: $E(\varepsilon_i) = 0$

وتعني هذه الفرضية أن الأخطاء لا تدخل في تفسير Y، إذ أنها تعبر عن حدود عشوائية تأخذ قيما سالبة، موجبة أو معدومة لا يمكن قياسها أو تحديدها بدقة، وتخضع لقوانين الاحتمال، بحيث يكون وسطها أو توقعها الرياضي مساويا للصفر: $E(\varepsilon_i) = 0, \forall i = 1.....n$

ب. الفرضية الثانية: تجانس (ثبات) تباين الأخطاء Homoscedasticity:

وهو ما يعني أن تشتتها حول المتوسط ثابت، ونعبر عنها رياضيا بالكتابة:

$$Var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2 \quad , \forall i = 1.....n$$

ج. الفرضية الثالثة: عدم وجود ارتباط ذاتي بين الأخطاء: بمعنى أن التباينات المشتركة لأخطاء الملاحظات المختلفة تكون معدومة، وهذا على مختلف مشاهدات مكونات العينة، ونعبر عنها رياضيا كما يلي:

$$\text{Cov}(\varepsilon_i, \varepsilon_j) = \text{E}(\varepsilon_i \varepsilon_j) = 0 \quad , \forall i \neq j \qquad i, j = 1.....n$$

د. الفرضية الرابعة: تتعلق بقيم المتغير المستقل X_i، تتمثل في أن المعطيات التي جمعت بالنسبة لهذا المتغير قادرة على إظهار تأثيرها في تغير المتغير التابع Y_i، بحيث تكون قيمة واحدة على الأقل مختلفة عن بقية القيم[1]،

1- عبد الحميد عبد المجيد البلداوي، 1997، ص 506.

أي مهما يكن حجم العينة n: يكون المقـدار $(1/n)\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\neq 0$ ،

أي أن الأخطاء تكون مستقلة عن X_i :

$$Cov(X_i,\varepsilon_i)=0, \forall i=1,...,n$$

2. تقدير معالم النموذج:

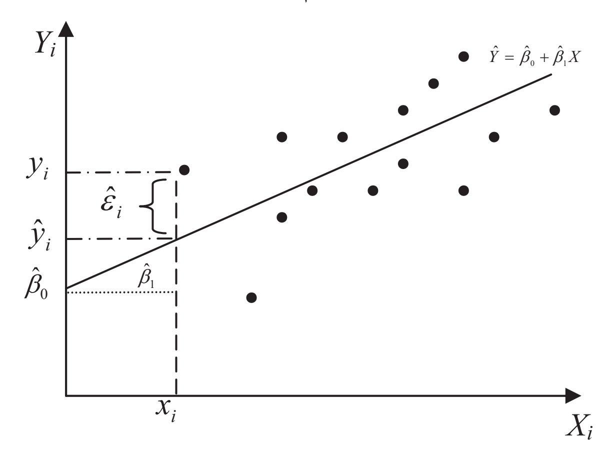
سنرمز فيما يلي إلى القيمة المقدرة لـ Y_i. (الحقيقية) بـ $\hat{Y}_i$.:

عند سحب عينة مكونة من n ثنائية: $(X_1,Y_1),........,(X_n,Y_n)$ يكون التسـاؤل حول الخط الذي يُعبِّر بكيفية جيدة عن العلاقة: $Y_i=\beta_0+\beta_1X_i+\varepsilon_i$ ولـذلك يجـب تقدير المعاملين β_0 و β_1. فعند تمثيل ثنائيات المشاهدات في بيان يُظهر لنا تشـتت هـذه المشاهدات (الشكل رقم (1))، يكون هدفنا هو البحث عن تصحيح يعبر تعبيرا جيدا عن العلاقة أعلاه. في هده الحالة نقترح طريقة المربعات الصغرى.

1.2. طريقة المربعات الصغرى:

إن هذه الطريقة تحاول إيجاد أحسن تصحيح خطي بتدنئه مربعات الانحـراف (بـين المشاهدات الفعلية والمقدرة) $\sum_{i=1}^{n}\hat{\varepsilon}_i^2$، حيث: $\hat{\varepsilon}_i=Y_i-\hat{Y}_i$. (أنظر الشكل رقم (1))[1].

الشكل رقم (1): الهدف من طريقة المربعات الصغرى



1- هتهات سعيد، 2006، ص. 99.

وهذا ما يمكن كتابته رياضيا بـ: $Min\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \underset{\hat{\beta}_0,\hat{\beta}_1}{Min}\sum_{i=1}^{n}\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right)^2$

والشرط اللازم لتدنئة هذه العلاقة هو أن تكون المشتقات الجزئية بالنسبة $\hat{\beta}_1, \hat{\beta}_0$ معدومة أي:

$$\begin{cases} \frac{\partial}{\partial\hat{\beta}_0}\sum_i \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right)^2 = 0 \\ \frac{\partial}{\partial\hat{\beta}_1}\sum_i \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right)^2 = 0 \end{cases}$$

بعد حل جملة المعادلين السابقة نتحصل على تقدير معلمتي النموذج:

$$\begin{cases} \hat{\beta}_1 = \dfrac{n\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{n\sum_i X_i^2 - \left(\sum_i X_i\right)^2} \\ \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X} \end{cases}$$

ومن المفيد استخدام صيغة مكافئة لتقدير $\hat{\beta}_1$: $\hat{\beta}_1 = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$

ويكون النموذج المقدر (خط الانحدار) بطريقة المربعات الصغرى المقدرة (*OLS*) كما يلي:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

2.2. خصائص مقدرات المربعات الصغرى:

1.2.2. خاصية عدم التحيز:

التحيز هو ذلك الفرق بين مقدرة ما ووسط توزيعها، فإذا كان هذا الفرق يختلف عن الصفر نقول عن ذلك المقدر بأنه متحيز. وإذا عدنا إلى مقدرتي المربعات الصغرى فإننا نجد $E\left(\hat{\beta}_1\right) = \beta_1 , E\left(\hat{\beta}_0\right) = \beta_0$ ومنه نقول أن $\hat{\beta}_0$ و $\hat{\beta}_1$ هما مقدرتين غير متحيزتين لـ β_0. و β_1 على التوالي.

البرهان: نبرهن أن $\hat{\beta}_0$ و $\hat{\beta}_1$ مقدران غير متحيزين.

- بالنسبة لـ $\hat{\beta}_1$:

لدينا المقدر $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$ حيث: $y_i = Y_i - \overline{Y}$ و $x_i = X_i - \overline{X}$

نضع: $w_i = \frac{x_i}{\sum_{i=1}^{n} x_i^2}$ ، فيمكن كتابة المقدر كما يلي: $\hat{\beta}_1 = \sum_{i=1}^{n} w_i y_i$ فهو خطي لأن:

$$\hat{\beta}_1 = \sum_{i=1}^{n} w_i y_i = w_1 y_1 + w_2 y_2 + ... + w_n y_n$$

خصائص ω_i :

- $\sum_{i=1}^{n} w_i = 0$
- $\sum_{i=1}^{n} w_i x_i = 1$
- $\sum_{i=1}^{n} w_i^2 = \frac{1}{\sum_{i=1}^{n} x_i^2}$

نعلم أن $\hat{\beta}_1$ يكتب على الشكل:

$\hat{\beta}_1 = \sum_{i=1}^{n} w_i [Y_i - \overline{Y}]$ حيث: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ و

$$\overline{Y} = \beta_0 + \beta_1 \overline{X} + \overline{\varepsilon}$$

وعليه: $\hat{\beta}_1 = \sum_{i=1}^{n} w_i [\beta_0 + \beta_1 X_i + \varepsilon_i - \beta_0 - \beta_1 \overline{X} - \overline{\varepsilon}]$

نحصل على:

$$\hat{\beta}_1 = \sum_{i=1}^{n} w_i [\beta_1 X_i + \varepsilon_i - \beta_1 \overline{X} - \overline{\varepsilon}] = \sum_{i=1}^{n} w_i [\beta_1 (X_i - \overline{X}) + (\varepsilon_i - \overline{\varepsilon})]$$

المقدار $\bar{\varepsilon} - \varepsilon_i$ يؤول إلى ε_i لأن $\bar{\varepsilon}$ مسـ ـاو للصـ ـفر بالأمـ ـل الرياضـ ـي و $\sum_{i=1}^{n} w_i x_i = 1$ ، إذن:

$$\hat{\beta}_1 = \beta_1 + \sum_{i=1}^{n} w_i \varepsilon_i \qquad \text{(} \hat{\beta}_1 \text{ عشوائي)}$$

بإدخال التوقع الرياضي على الطرفين: $E(\hat{\beta}_1) = E(\beta_1) + \sum_{i=1}^{n} w_i E(\varepsilon_i)$ و

باعتبار أن: $E(\beta_1) = \beta_1$ (β_1 معلم نظري غير عشوائي) و $E(\varepsilon_i) = 0$ ، ومنه نستنتج أن $E(\hat{\beta}_1) = \beta_1$ فهو مقدر غير متحيز. نذكر فقط أن مقدار التحيز $\hat{\beta}_1 - \beta_1$ يساوي إلى $\sum_{i=1}^{n} w_i \varepsilon_i$.

- بالنسبة لـ $\hat{\beta}_0$:

نعلم أن: $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ حيث: $\bar{Y} = \beta_0 + \beta_1 \bar{X} + \bar{\varepsilon}$

$\hat{\beta}_0 = \beta_0 + \beta_1 \bar{X} + \bar{\varepsilon} - \hat{\beta}_1 \bar{X}$ و بإدخال التوقع الرياضي على الطرفين، نحصل على:

$E(\hat{\beta}_0) = \beta_0 + \beta_1 \bar{X} + \frac{1}{n} \sum_{i=1}^{n} E(\varepsilon_i) - E(\hat{\beta}_1) \bar{X}$ مع العلم أن $E(\hat{\beta}_1) = \beta_1$

نستنتج أن $E(\hat{\beta}_0) = \beta_0$. المقدر $\hat{\beta}_0$ غير متحيز.

يمكن أيضا كتابة الثابتة المقدرة $\hat{\beta}_0$ على الشكل التالي:

$$\hat{\beta}_0 = \frac{1}{n} \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} w_i y_i \bar{X} = \frac{1}{n} \sum_{i=1}^{n} Y_i - \bar{X} \sum_{i=1}^{n} w_i (Y_i - \bar{Y}) = \frac{1}{n} \sum_{i=1}^{n} Y_i - \bar{X} \sum_{i=1}^{n} w_i Y_i$$

ومنه: $\hat{\beta}_0 = \sum_{i=1}^{n} \left(\frac{1}{n} - w_i \bar{X} \right) Y_i$

نضع: $w_i^* = \frac{1}{n} - w_i \bar{X}$

المقدر يكتب: $\hat{\beta}_0 = \sum_{i=1}^{n} w_i^* Y_i = w_1^* Y_1 + w_2^* Y_2 + \ldots.. + w_n^* Y_n$ فهو إذن خطي.

2.2.2. أفضل مقدر خطي غير متحيز BLUE ومتسق:

تنطلق هذه الفكرة من نظرية Gauss-Markov والتي تقول" من بين المقدرات الخطية وغير المتحيزة، تكون مقدرتا المربعات الصغرى العادية $\hat{\beta}_0$ و $\hat{\beta}_1$ أفضل مقدرتين خطيتين وغير متحيزتين، حيث أن لها أصغر تباين ممكن مقارنة مع بقية المقدرات الخطية وغير المتحيزة الأخرى".

إذا واجهنا مشكلة تحيز مقدرة ما، فإننا ننظر إلى الخاصية التقاربية لذلك المقدر، ويحدث ذلك لما يكون المتغير المستقل X_i عبارة عن متغير تابع ومبطأ بفترة زمنية ما، ونقول عن $\hat{\beta}_1$ بأنه مقدر متسق (Consistent Estimator)، إذا كان: كلما $n \rightarrow \infty$ فإن توزيع المعاينة لـ $\hat{\beta}_1$. يقترب من القيمة الحقيقية β_1، ونقول أن النهاية الاحتمالية للمقدر $\hat{\beta}_1$ هي β_1 ونكتب: $\underset{n\rightarrow\infty}{p\lim}(\hat{\beta}_1) = \beta_1$

لكن هذا الشرط غير كاف للحصول على مقدر متسق، بل يجب أن تكـون قيمتـا التحيز والتباين تقتربان أو تساويان الصفر كلما اقترب n من ما لا نهاية أي:

$$\lim_{n\rightarrow\infty} E(\hat{\beta}_1) = \underset{n\rightarrow\infty}{p\lim}(\hat{\beta}_1) = \beta_1$$
$$\lim_{n\rightarrow\infty} \text{var}(\hat{\beta}_1) = \underset{n\rightarrow\infty}{p\lim}\, \text{var}(\hat{\beta}_1) = 0$$

وبتحقق هذين الشرطين، نقول عن المقدر $\hat{\beta}_1$ بأنه مقدر متسق للمعلمة الحقيقيـة. إن المقدرات المتحصل عليها لكل من β_0، β_1 و σ^2 سواء بطريقة المربعات الصغرى أو غيرها هي تقديرات نقطية، ولكن من المهم أن يكون لدى الاقتصادي أكثر من اختيار، ولذلك يجب أن نبني مجالا لهذه المقدرات وذلك بقبول مستوى ثقة معين وهو ما نسميه بالتقـدير المجالي للمعالم.

مثال 1:

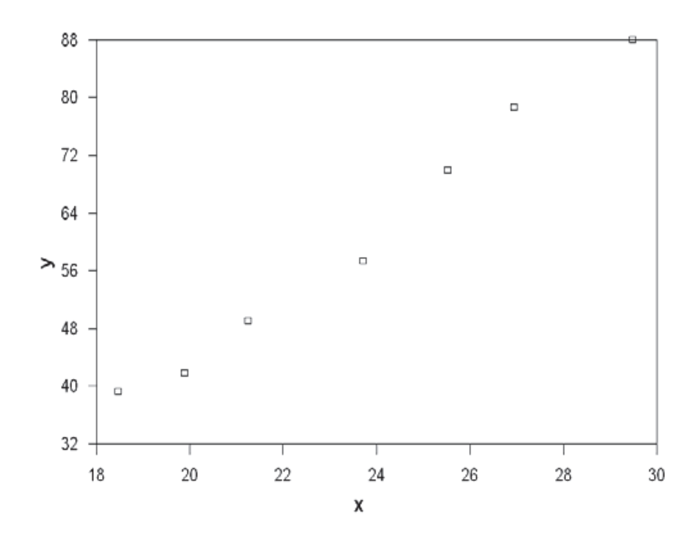
يبين الجدول التالي تطور كل من الاستهلاك و الدخل المتاح في الجزائر خلال السنوات 2001 و 2007:

الجدول (1): تطور الاستهلاك و الدخل خلال 10 سنوات

الدخل المتاح X_i	الاستهلاك الإجمالي Y_i	السنة i
39.25	18.47	2001
41.84	19.89	2002
49.06	21.26	2003
57.31	23.71	2004
69.89	25.53	2005
78.64	26.95	2006
88.00	29.48	2007

يظهر الشكل (1-2) الزوج الخاص بمعطيات الاستهلاك و الدخل. نلاحظ أن العلاقة التي تربط بينهما خطية و يظهر جليا من خلال سحاب النقاط.

الشكل رقم (2): سحاب النقاط للزوج (الاستهلاك–الدخل)



انطلاقا من معطيات الجدول (1)، نقوم بحساب كل من $\hat{\beta}_0$ و $\hat{\beta}_1$.

الجدول (2): حساب معاملات الانحدار

السنة i	Y_i	X_i	$Y_i - \bar{Y}$	$X_i - \bar{X}$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$
2001	18.47	39.25	-5.14	-21.32	109.58	454.54	26.41
2002	19.89	41.84	-3.72	-18.73	69.67	350.81	13.83
2003	21.26	49.06	-2.35	-11.51	27.04	132.48	5.52
2004	23.71	57.31	0.10	-3.26	-0.32	10.62	0.01
2005	25.53	69.89	1.92	9.32	17.89	86.86	3.68
2006	26.95	78.64	3.34	18.07	60.35	326.52	11.15
2007	29.48	88.00	5.87	27.43	161.01	752.40	34.45
المجموع	165.29	423.99	0	0	445.24	2114.25	95.05
المتوسط	23.61	60.57	0	0	63.60	302.03	13.57

لدينا:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{7}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{7}(X_i - \bar{X})^2} = \frac{445.24}{2114.25} = 0.21$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X} = 23.61 - (0.21)(60.57) = 10.89$$

وعليه النموذج المقدر يكتب كما يلي:

$$\hat{Y}_i = 0.21X_i + 10.89$$

3. توزيع المعاينة للمقدرات و التقدير المجالي للمعالم:

1.3. حساب تباينات المقدرات:

لبناء مجال الثقة للمعالم، يتعين معرفة تباين كل من $\hat{\beta}_0$، $\hat{\beta}_1$ و البواقي.

- تباين : $\hat{\beta}_0$

نحسب المقدار : $\hat{\beta}_0 - \beta_0$

$$\hat{\beta}_0-\beta_0=\beta_1\overline{X}+\overline{\varepsilon}-\hat{\beta}_1\overline{X}=\overline{\varepsilon}-(\hat{\beta}_1-\beta_1)\overline{X}=\sum_{i=1}^{n}\left(\frac{1}{n}-w_i\overline{X}\right)\varepsilon_i=\sum_{i=1}^{n}w_i^*\varepsilon_i$$

وعليه:

$$\left(\hat{\beta}_0-\beta_0\right)^2=\sum_{i=1}^{n}w_i^{*2}\varepsilon_i^2+2\sum_i\sum_j w_i^*w_j^*\varepsilon_i\varepsilon_j$$

بإدخال التوقع الرياضي على الطرفين:

$$E\left(\left(\hat{\beta}_0-\beta_0\right)^2\right)=\sum_{i=1}^{n}w_i^{*2}E(\varepsilon_i^2)+2\sum_i\sum_j w_i^*w_j^*E(\varepsilon_i\varepsilon_j)$$

بما أن $E(\varepsilon_i\varepsilon_j)=0$ ، $E(\varepsilon_i^2)=\sigma_\varepsilon^2$ و

$$\sum_{i=1}^{n}w_i^{*2}=\sum_{i=1}^{n}\left(\frac{1}{n^2}+w_i^2\overline{X}^2-\frac{2}{n}w_i\overline{X}\right)=\frac{1}{n}+\sum_{i=1}^{n}w_i^2\overline{X}^2$$

فإن:

$$\operatorname{var}(\hat{\beta}_0)=\left(\frac{1}{n}+\frac{\overline{X}^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2}\right)\sigma_\varepsilon^2=\left(\frac{\sum_{i=1}^{n}X_i^2}{n\sum_{i=1}^{n}(X_i-\overline{X})^2}\right)\sigma_\varepsilon^2$$

- <u>تباين $\hat{\beta}_1$:</u>

نذكر أن المقدار $\hat{\beta}_1-\beta_1$ يكتب على الشكل التالي:

$$\hat{\beta}_1-\beta_1=\sum_{i=1}^{n}w_i\varepsilon_i$$

وعليه:

$$\left(\hat{\beta}_1-\beta_1\right)^2=\left(\sum_{i=1}^{n}w_i\varepsilon_i\right)^2=\sum_{i=1}^{n}w_i^2\varepsilon_i^2+2\sum_i\sum_j w_iw_j\varepsilon_i\varepsilon_j$$

بإدخال التوقع الرياضي على الطرفين:

$$E\left(\left(\hat{\beta}_1-\beta_1\right)^2\right)=\sum_{i=1}^{n}w_i^2E(\varepsilon_i^2)+2\sum_i\sum_j w_iw_jE(\varepsilon_i\varepsilon_j)$$

بما أن $E(\varepsilon_i\varepsilon_j)=0$ و $E(\varepsilon_i^2)=\sigma_\varepsilon^2$ فإن:

$$\operatorname{var}(\hat{\beta}_1)=\sigma_\varepsilon^2\sum_{i=1}^{n}w_i^2=\frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

العلاقة بين $\operatorname{var}(\hat{\beta}_1)$ و $\operatorname{var}(\hat{\beta}_0)$ تعطى بالعلاقة التالية:

$$\operatorname{var}(\hat{\beta}_0)=\frac{\sigma_\varepsilon^2}{n}+\overline{X}^2\operatorname{var}(\hat{\beta}_1)$$

وبناءا على هذا التعريف تكون الانحرافات المعيارية (Standard déviations) هي الجذور التربيعية لتباينات المقدرات، أما الأخطاء المعيارية (Standard errors) فهي الجذور التربيعية لمقدرات الانحرافات المعيارية أي:

$$\hat{\sigma}_{\hat{\beta}_0} = \sqrt{\mathrm{var}\left(\hat{\beta}_0\right)} = \sigma_\varepsilon \sqrt{\frac{\sum X_i^2}{n\sum (X_i - \overline{X})^2}}$$

$$\hat{\sigma}_{\hat{\beta}_1} = \sqrt{\mathrm{var}\left(\hat{\beta}_1\right)} = \frac{\sigma_\varepsilon}{\sqrt{\sum (X_i - \overline{X})^2}}$$

نلاحظ أن تباين كل مقدر غير معروف لأنه يرتبط بتباين الأخطاء النظري σ_ε^2، فينبغي في هده الحالة تقدير تباين الأخطاء للحصول على تباين البواقي:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1.........n$$

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

$$\overline{Y} = \beta_0 + \beta_1 \overline{X} + \overline{\varepsilon}$$

$$\hat{\varepsilon}_i = Y_i - \hat{Y}_i = (\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1) X_i + \varepsilon_i$$

ولكن المقدار $\beta_0 - \hat{\beta}_0$ يمكن كتابته بعد تعويض $\hat{\beta}_0$ بقيمته كالتالي:

$$\beta_0 - \hat{\beta}_0 = \beta_0 - \overline{Y} + \hat{\beta}_1 \overline{X} = \beta_0 - (\beta_0 + \beta_1 \overline{X} + \overline{\varepsilon}) + \hat{\beta}_1 \overline{X} = (\hat{\beta}_1 - \beta_1)\overline{X} - \overline{\varepsilon}$$

نكتب المعادلة كما يلي:

$$\hat{\varepsilon}_i = (\beta_1 - \hat{\beta}_1) X_i + (\hat{\beta}_1 - \beta_1)\overline{X} + (\varepsilon_i - \overline{\varepsilon}) = (\beta_1 - \hat{\beta}_1)(X_i - \overline{X}) + (\varepsilon_i - \overline{\varepsilon})$$

بتربيع طرفي المعادلة، نحصل على:

$$\hat{\varepsilon}_i^2 = (\beta_1 - \hat{\beta}_1)^2 (X_i - \overline{X})^2 + (\varepsilon_i - \overline{\varepsilon})^2 + 2(\beta_1 - \hat{\beta}_1)(X_i - \overline{X})(\varepsilon_i - \overline{\varepsilon})$$

ولدينا سابقا: $\hat{\beta}_1 - \beta_1 = \sum_{i=1}^{n} w_i \varepsilon_i \Rightarrow \left(\hat{\beta}_1 - \beta_1\right)^2 = \left(\sum_{i=1}^{n} w_i \varepsilon_i\right)^2$

وللتسهيل، نضع: $(X_i - \overline{X})^2 = x_i^2$

$$\hat{\varepsilon}_i^2 = (\sum_{i=1}^{n} w_i \varepsilon_i)^2 x_i^2 + (\varepsilon_i - \overline{\varepsilon})^2 - 2(\sum_{i=1}^{n} w_i \varepsilon_i) x_i (\varepsilon_i - \overline{\varepsilon})$$

بإدخال التوقع الرياضي على الطرفين:

$$E(\hat{\varepsilon}_i^2) = E\left[(\sum_{i=1}^{n} w_i \varepsilon_i)^2 x_i^2 + (\varepsilon_i - \bar{\varepsilon})^2 - 2(\sum_{i=1}^{n} w_i \varepsilon_i) x_i (\varepsilon_i - \bar{\varepsilon}) \right]$$

نبحث عن توقع هدا المقدار بمراحل ودلك بإيجاد توقع كل حد على حدة:

<u>الحد الأول:</u>

$$\left(\sum_{i=1}^{n} w_i \varepsilon_i \right)^2 = \sum_{i=1}^{n} w_i^2 \varepsilon_i^2 + 2\sum_i \sum_j w_i w_j \varepsilon_i \varepsilon_j$$

بإدخال التوقع الرياضي على الطرفين:

$$E\left(\left(\sum_{i=1}^{n} w_i \varepsilon_i \right)^2 \right) = \sum_{i=1}^{n} w_i^2 E(\varepsilon_i^2) + 2\sum_i \sum_j w_i w_j E(\varepsilon_i \varepsilon_j) = \sigma_\varepsilon^2 \sum_{i=1}^{n} w_i^2$$

$$E\left(\left(\sum_{i=1}^{n} w_i \varepsilon_i \right)^2 \right) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n} x_i^2}$$

<u>الحد الثاني:</u>

$$\varepsilon_i - \bar{\varepsilon} = \varepsilon_i - \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i = \varepsilon_i - \frac{1}{n} (\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n)$$

$$(\varepsilon_i - \bar{\varepsilon})^2 = \varepsilon_i^2 + \frac{1}{n^2} (\varepsilon_1^2 + \ldots + \varepsilon_n^2 + 2\sum_i \sum_j \varepsilon_i \varepsilon_j) - \frac{2}{n} \varepsilon_i (\varepsilon_1 + \ldots + \varepsilon_n)$$

بإدخال التوقع الرياضي على الطرفين:

$$E((\varepsilon_i - \bar{\varepsilon})^2) = E(\varepsilon_i^2) + \frac{1}{n^2} E(\sum_{i=1}^{n} \varepsilon_i^2) + 2\sum_i \sum_j E(\varepsilon_i \varepsilon_j) - \frac{2}{n} \sum E(\varepsilon_i \varepsilon_j)$$

$$i \neq j \;,\; 2\sum_i \sum_j E(\varepsilon_i \varepsilon_j) = 0 \,,\; \frac{1}{n^2} E(\sum_i \varepsilon_i^2) = \sigma^2 / n \,,\; E(\varepsilon_i^2) = \sigma^2$$

$$\frac{1}{n} \sum E(\varepsilon_i \varepsilon_j) = \sigma^2 / n$$ (لما $i = j$)

إذن:

$$E((\varepsilon_i - \bar{\varepsilon})^2) = \sigma^2 + \frac{\sigma^2}{n} - 2\frac{\sigma^2}{n}$$

الحد الثالث:

لدينا: $$(\sum_{i=1}^{n} w_i\varepsilon_i)(\varepsilon_i-\bar{\varepsilon})=(\sum_{i=1}^{n} w_i\varepsilon_i)\varepsilon_i-(\sum_{i=1}^{n} w_i\varepsilon_i)\bar{\varepsilon}$$

بإدخال التوقع الرياضي:

$$E\left((\sum_{i=1}^{n} w_i\varepsilon_i)(\varepsilon_i-\bar{\varepsilon})\right)=E\left((\sum_{i=1}^{n} w_i\varepsilon_i)\varepsilon_i\right)-E\left((\sum_{i=1}^{n} w_i\varepsilon_i)\bar{\varepsilon}\right)$$

- $E\left((\sum_{i=1}^{n} w_i\varepsilon_i)\varepsilon_i\right)=w_i\sigma^2$

$$E\left((\sum_{j} w_j\varepsilon_j)\varepsilon_i\right)=\sum_{j} w_j E(\varepsilon_i\varepsilon_j)$$

$$E(\varepsilon_i\varepsilon_j)=\begin{cases}0, i\neq j\\ \sigma^2, i=j\end{cases}$$

ومنه: $$\sum_{j} w_j E(\varepsilon_i\varepsilon_j)=\sum_{i} w_i\sigma^2$$

- $(\sum_{i=1}^{n} w_i\varepsilon_i)\bar{\varepsilon}=(\sum_{i} w_i\varepsilon_i)\frac{1}{n}\sum_{i}\varepsilon_i=(\sum_{i} w_i\varepsilon_i)\frac{\varepsilon_1}{n}+(\sum_{i} w_i\varepsilon_i)\frac{\varepsilon_2}{n}+\ldots+(\sum_{i} w_i\varepsilon_i)\frac{\varepsilon_n}{n}$

$$=\frac{w_1\sigma^2}{n}+\frac{w_2\sigma^2}{n}+\ldots+\frac{w_n\sigma^2}{n}$$

وعليه: $$E\left((\sum_{i=1}^{n} w_i\varepsilon_i)(\varepsilon_i-\bar{\varepsilon})x_i\right)=\sum_{i} x_i w_i\sigma^2$$

بجمع الحد الأول، الثاني و الثالث، يكون لدينا:

$$\sum_{i=1}^{n} E(\hat{\varepsilon}_i^2)=\frac{\sigma^2}{\sum x_i^2}\sum x_i^2+n\sigma^2+\sigma^2+(-2\sigma^2)-2\sigma^2\sum_{i} x_i w_i$$

$$=n\sigma^2-2\sigma^2=(n-2)\sigma^2$$

نلاحظ أن:
$$E(\sum_{i=1}^{n}\hat{\varepsilon}_i^2) = (n-2)\sigma^2 \Rightarrow E\left(\frac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{n-2}\right) = \sigma^2$$

إذن: $\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{n-2}$ و هو تقدير غير متحيز

القيمة التي في البسط تعبر عن مجموع مربعات البواقي حيث $\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$ أما $n-2$ فهي درجة الحرية، تعبر عن حجم العينة ناقص 2 و ذلك لوجود معلمين للتقدير في النموذج.

2.3. بناء مجال الثقة للمعالم:

بمعرفة توزيع $\hat{\beta}_0$ و $\hat{\beta}_1$ يمكن تكوين مجالات ثقة وإجراء اختبار الفرضيات الموضوعة حول معالم الانحدار β_0 و β_1 على التوالي، نعطي مجالا للقيم التي يمكن أن تحتوي عليها معالم الانحدار الحقيقية، مع كل مجال ثقة نضع مستوى إحصائيا للمعنوية، حيث أن احتمال احتواء المجال المذكور على معلمة الانحدار الحقيقية يكون واحد مطروحا منه مستوى المعنوية، أي $(1-\alpha)$، ولتكوين مجال الثقة من التوزيع t بالنسبة للمعلمين β_0 و β_1 نكتب القانون الخاص لكل معلمة:

في حالة $n \leq 30$ و σ^2 غير معروف:

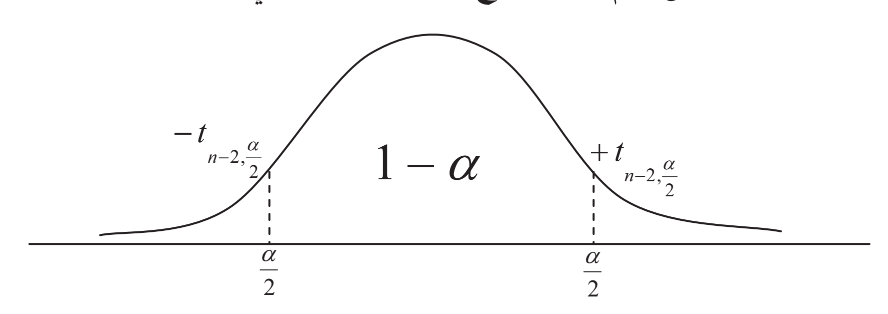
$$\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \sim t_{(n-2)}$$
$$\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \sim t_{(n-2)}$$

عند مستوى معنوية $(\alpha\%)$ يكون مجال الثقة لكلا المعلمين:

$$\Pr\left[-t_{n-2,\frac{\alpha}{2}} \le \frac{\hat{\beta}_0-\beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \le +t_{n-2,\frac{\alpha}{2}}\right]=1-\alpha$$

$$\Pr\left[-t_{n-2,\frac{\alpha}{2}} \le \frac{\hat{\beta}_1-\beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \le +t_{n-2,\frac{\alpha}{2}}\right]=1-\alpha$$

الشكل رقم (3): توزيع المعاينة لـ . $\hat{\beta}_1$ ثنائي الطرف



إذا ضربنا (داخل الاحتمال) كل الأطراف بواسطة $\hat{\sigma}_{\hat{\beta}_0}$ ($\hat{\sigma}_{\hat{\beta}_1}$) وأضفنا β_0 (β_1) لأطراف المتراجحة نجد:

$$\beta_0 \in \left[\hat{\beta}_0 - t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}, \hat{\beta}_0 + t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}\right]$$

$$\beta_1 \in \left[\hat{\beta}_1 - t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}\right]$$

$t_{n-2,\frac{\alpha}{2}}$: القيمة الحرجة لتوزيع *Student* بدرجة حرية $n-2$ و نسبة معنوية $(\alpha\%)$ ونجد من جدول التوزيع القيمة المحسوبة.

في حالة $n > 30$ و σ^2 معروف:

$$\hat{\beta}_1 \longrightarrow N\left(\beta_1, \frac{\hat{\sigma}_\varepsilon^2}{\sum_i x_i^2}\right)$$

$$\hat{\beta}_0 \longrightarrow N\left(\beta_0, \hat{\sigma}_\varepsilon^2 \frac{\sum_i X_i^2}{n\sum_i x_i^2}\right)$$

$$\frac{\hat{\beta}_0-\beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \longrightarrow N(0,1)$$

$$\frac{\hat{\beta}_1-\beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \longrightarrow N(0,1)$$

عند مستوى معنوية $(\alpha\%)$ يكون مجال الثقة لكلا المعلمين:

$$\Pr\left[-z_{\frac{\alpha}{2}} \leq \frac{\hat{\beta}_0-\beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \leq +z_{\frac{\alpha}{2}}\right]=1-\alpha$$

$$\Pr\left[-z_{\frac{\alpha}{2}} \leq \frac{\hat{\beta}_1-\beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \leq +z_{\frac{\alpha}{2}}\right]=1-\alpha$$

نفس الشيء، نضرب (داخل الاحتمال) كل الأطراف بواسطة $\hat{\sigma}_{\hat{\beta}_0}$ ($\hat{\sigma}_{\hat{\beta}_1}$) ونضيف β_0 (β_1) لأطراف المتراجحة نجد:

$$\beta_0 \in \left[\hat{\beta}_0 - z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}, \hat{\beta}_0 + z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}\right]$$

$$\beta_1 \in \left[\hat{\beta}_1 - z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}\right]$$

$z_{\frac{\alpha}{2}}$: القيمة الحرجة للتوزيع الطبيعي بنسبة معنوية $(\alpha\%)$ ونجد من جدول التوزيع القيمة المحسوبة. كلما كان مجال الثقة ضيقا كلما كان المقدر أحسن، لأن الأخطاء المعيارية تكون أصغر.

نبني أيضا مجال الثقة لـ σ^2. لدينا:

$$\chi^2_\alpha(n-2) \sim \frac{(n-2)\hat{\sigma}^2_\varepsilon}{\sigma^2_\varepsilon}$$

$\chi^2_\alpha(n-2)$: القيمة الحرجة لتوزيع χ^2 بدرجة حرية $n-2$. يكون مجال الثقة:

$$\Pr\left[\chi^2_{\alpha/2} \leq \frac{(n-2)\hat{\sigma}^2_\varepsilon}{\sigma^2_\varepsilon} \leq \chi^2_{1-\alpha/2}\right]=1-\alpha \Rightarrow \Pr\left[\frac{(n-2)\hat{\sigma}^2_\varepsilon}{\chi^2_{1-\alpha/2}} \leq \sigma^2_\varepsilon \leq \frac{(n-2)\hat{\sigma}^2_\varepsilon}{\chi^2_{\alpha/2}}\right]=1-\alpha$$

يكون مجال الثقة لتباين الأخطاء:

$$\sigma_{\varepsilon}^{2} \in \left[\frac{(n-2)\hat{\sigma}_{\varepsilon}^{2}}{\chi_{1-\alpha/2}^{2}}, \frac{(n-2)\hat{\sigma}_{\varepsilon}^{2}}{\chi_{\alpha/2}^{2}} \right]$$

مثال 2:

نأخذ معطيات المثال الأول و نقوم بحساب تباين كل من $\hat{\beta}_0$ و $\hat{\beta}_1$ و الأخطاء المعيارية ولكن بعد حساب تباين البواقي $\hat{\sigma}_{\varepsilon}^{2}$ و من تم نبني مجال الثقة لكل معلم بنسبة ثقة 95%. يجب أولا حساب القيم المقدرة $\hat{Y}_i$ و بواقي التقدير $\hat{\varepsilon}_i$ كما هو مبين في الجدول (3).

الجدول (3): حساب بواقي التقدير

$\hat{\varepsilon}_i^2$	$\hat{\varepsilon}_i$	$\hat{Y}_i$	Y_i	i
0.42	-0.65	19.12	18.47	1
0.04	0.22	19.66	19.89	2
0.005	0.07	21.18	21.26	3
0.61	0.78	22.92	23.71	4
0.002	-0.04	25.57	25.53	5
0.21	-0.46	27.41	26.95	6
0.008	0.09	29.38	29.48	7
1.32	0			المجموع

حيث تم حساب قيم $\hat{Y}_i$ انطلاقا من الانحدار الخطي $\hat{Y}_i = 0.21X_i + 10.89$ و بواقي التقدير $\hat{\varepsilon}_i$ من المعادلة $\hat{\varepsilon}_i = Y_i - \hat{Y}_i$

نحسب تباين البواقي:

$$\hat{\sigma}_{\varepsilon}^{2} = \frac{\sum_{i=1}^{7}\hat{\varepsilon}_i^2}{n-2} = \frac{1.32}{7-2} = 0.26$$

و الذي يسمح بحساب تباين كل مقدر:

$$\hat{\sigma}_{\hat{\beta}_1}^{2} = \mathrm{var}(\hat{\beta}_1) = \frac{\hat{\sigma}_{\varepsilon}^{2}}{\sum_{i=1}^{7}(X_i - \overline{X})^2} = \frac{0.26}{2114.25} = 0.000124$$

$$\hat{\sigma}_{\hat{\beta}_0}^{2} = \mathrm{var}(\hat{\beta}_0) = \frac{\hat{\sigma}_{\varepsilon}^{2}}{n} + \overline{X}^2\,\mathrm{var}(\hat{\beta}_1) = \frac{0.26}{7} + (60.57)^2(0.000124) = 0.49$$

و الأخطاء المعيارية:

$$\hat{\sigma}_{\hat{\beta}_1} = \sqrt{\operatorname{var}(\hat{\beta}_1)} = \sqrt{0.000124} = 0.011$$

$$\hat{\sigma}_{\hat{\beta}_0} = \sqrt{\operatorname{var}(\hat{\beta}_0)} = \sqrt{0.49} = 0.7$$

يمكن بناء مجالات ثقة للمعالم:

$$\beta_0 \in \left[\hat{\beta}_0 - t_{0.025}\hat{\sigma}_{\hat{\beta}_0}, \hat{\beta}_0 + t_{0.025}\hat{\sigma}_{\hat{\beta}_0}\right]$$

$$\beta_1 \in \left[\hat{\beta}_1 - t_{0.025}\hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + t_{0.025}\hat{\sigma}_{\hat{\beta}_1}\right]$$

حيث $t_{0.025}$ هي القيمة الحرجة لتوزيع ستيودنت بنسبة معنوية 5% و درجة حرية $n-2=5$ والتي تعادل القيمة 2.570 في جدول توزيع ستيودنت. بالتطبيق العددي لدينا:

$$\beta_0 \in [10.89 - 2.570 \times 0.7 , \; 10.89 + 2.570 \times 0.7]$$

أي:

$$\beta_0 \in [9.09 \; , \; 12.68]$$

و:

$$\beta_1 \in [0.21 - 2.570 \times 0.011 \; , \; 0.21 + 2.57 \times 0.011]$$

هذا يعني:

$$\beta_1 \in [0.18 , 0.23]$$

4. التقدير بطريقة المعقولية العظمى

نفترض أن الأخطاء تتوزع توزيعا طبيعيا و النموذج دائما هو $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \; i = 1,...,n$. دالة كثافة الأخطاء تكتب على الشكل التالي:

$$f(\varepsilon_i) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right)$$

نسمي $L(\varepsilon_1,...,\varepsilon_n)$ دالة المعقولية العظمى، حيث:

$$L(\varepsilon_1,...,\varepsilon_n) = \prod_{i=1}^{n} f(\varepsilon_i) = \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right)$$

نعلم أن: $\varepsilon_i = Y_i - \beta_0 - \beta_1 X_i$ وعليه:

$$L(\varepsilon_1,...,\varepsilon_n) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\frac{\varepsilon_i^2}{\sigma^2}\right) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\frac{(Y_i - \beta_0 - \beta_1 X_i)^2}{\sigma^2}\right)$$

نقوم بتقدير معالم النموذج وذلك بتعظيم لوغاريتم دالة المعقولية العظمى، حيث:

$$\max_{\hat{\beta}_0,\hat{\beta}_1,\hat{\sigma}^2} \log L = \max_{\hat{\beta}_0,\hat{\beta}_1,\hat{\sigma}^2} \left\{ -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(Y_i - \beta_0 - \beta_1 X_i\right)^2 \right\}$$

نبحث عن شروط التعظيم حيث أن الشروط اللازمة هي:

$$\begin{cases} \dfrac{\partial \log L}{\partial \hat{\beta}_0} = 0 \Leftrightarrow -\dfrac{2}{\hat{\sigma}^2}\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)(-1) = 0 \\ \dfrac{\partial \log L}{\partial \hat{\beta}_1} = 0 \Leftrightarrow -\dfrac{2}{\hat{\sigma}^2}\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)(-X_i) = 0 \\ \dfrac{\partial \log L}{\partial \hat{\sigma}^2} = 0 \Leftrightarrow -\dfrac{n}{2}.\dfrac{1}{\hat{\sigma}^2}.\dfrac{1}{2\hat{\sigma}^4}\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2 = 0 \end{cases}$$

من المعادلات الثلاث يمكن بكل سهولة استخراج قيم المقدرات:

من المعادلة الأولى: $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$

من المعادلة الثانية: $$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

ومن المعادلة الثالثة للجملة: $$-n\hat{\sigma}^2 = -\sum_{i=1}^{n}\hat{\varepsilon}_i^2 \Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{n}$$

لنتذكر تقدير تباين الأخطاء بطريقة المربعات الصغرى فبرهننا أنه غير متحيز و لكن المقدر المتحصل عليه بطريقة المعقولية العظمى متحيز. فلنسم $\hat{\sigma}^2_{OLS}$ مقدر تباين الأخطاء بطريقة المربعات الصغرى و $\hat{\sigma}^2_{MLE}$ المقدر المتحصل عليه بطريقة المعقولية العظمى، لدينا:

$$\hat{\sigma}^2_{MLE} = \frac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{n} = \frac{n-2}{n}.\frac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{n-2}$$

بإدخال التوقع الرياضي على الطرفين:

$$E(\hat{\sigma}^2_{MLE}) = \left(\frac{n-2}{n}\right).E\left(\frac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{n-2}\right) = \left(\frac{n-2}{n}\right).E\left(\hat{\sigma}^2_{OLS}\right)$$

بما أن $\hat{\sigma}^2_{OLS}$ غير متحيز أي: $E(\hat{\sigma}^2_{OLS}) = \sigma^2$، إذن:

$E(\hat{\sigma}^2_{MLE}) = \left(\frac{n-2}{n}\right).\sigma^2 \Rightarrow E(\hat{\sigma}^2_{MLE}) \neq .\sigma^2$ وعليه نستنتج أن المقدر بطريقة المعقولية العظمى متحيز.

يكون $\hat{\sigma}^2_{MLE}$ أحسن تقدير لـ σ^2. إذا كان حجم العينة كبيرا أي أن $\frac{n-2}{n} \to 1$ لما $n \to \infty$ نستنتج أن: $E(\hat{\sigma}^2_{MLE}) = \sigma^2$ تسمى هده الخاصية بخاصية عدم التحيز التقاربي.

5. تحليل التباين و القدرة التفسيرية للنموذج

تساعد البواقي $\hat{\varepsilon}_i$ على قياس مدى تمثيل المعادلة المفروضة في النموذج لمشاهدات العينة، حيث أن القيمة الكبيرة للبواقي تعني بأن التمثيل يكون غير جيد والقيمة الصغيرة لها تعني تمثيلا جيدا للنموذج، إن المشكلة في استعمال البواقي كمقياس لجودة التوفيق هو أن قيمة البواقي تعتمد على المتغير التابع Y_i، الذي نعرفه حول وسطه انطلاقا من الشكل رقم (1) كما يلي:

$$Y_i = \hat{Y}_i + \hat{\varepsilon}_i$$
$$Y_i - \overline{Y} = \hat{Y}_i - \overline{Y} + \hat{\varepsilon}_i$$

وبتربيع طرفي المعادلة أعلاه وجمعها بالنسبة لكل i نجد:

$$\sum_i (Y_i - \overline{Y})^2 = \sum_i (\hat{Y}_i - \overline{Y})^2 + \sum_i \hat{\varepsilon}_i^2$$

وتعد هذه المعادلة مفيدة جدا لخدمة أغراضنا فيما يتعلق بقياس القدرة التفسيرية، ولذا من المهم أن نفحص بعناية معنى كل حد من حدودها[1]:

- ❖ $\sum_i (Y_i - \overline{Y})^2$ هو مجموع مربعات الانحرافات الكلية في المتغير Y: Total Sum of Squares (TSS)
- ❖ $\sum_i (\hat{Y}_i - \overline{Y})^2$ فهو مجموع مربعات الانحرافات المشروحة : Explained Sum of Squares (ESS)

1- المرسي السيد الحجازي، عبد القادر محمد عطية، 2001، ص 112.

❖ ويبقى الحد الأخير $\sum_i \hat{\varepsilon}_i^2$ الذي هو مجموع مربعات البواقي: Residual Sum of Squares (RSS)

نعيد صياغة المعادلة السابقة على $\boxed{TSS = ESS + RSS}$

وبتقسيم كل الأطراف على الانحرافات الكلية TSS نجد:

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS}$$

وعليه نعرف معامل التحديد $R^2 = r^2$ كما يلي[1]:

$$R^2 = r^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

معامل التحديد R^2 يقيس ويشرح نسبة الانحرافات الكلية أو التغيرات التي تحدث في المتغير التابع Y_i، والمشروحة بواسطة تغيرات المتغير المستقل X_i فهي نسبة تـ ـأثير المـ ـتغير المستقل على المتغير التابع، فهو إذن مقياس للقدرة التفسيرية للنموذج أي يخـ ـبر جـ ـودة التوفيق و الارتباط.

ويمكن حساب R^2 كالأتي:

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = 1 - \frac{\sum \hat{\varepsilon}_i^2}{\sum(Y_i - \bar{Y})^2}$$

ويعتبر R^2 من أهم المعاملات التي تقيس علاقة الارتباط بين متغيرين ووجود مثل هذه العلاقة يعني ضمنيا أن أحد هذين المتغيرين يعتمد في تغيره أو في حدوثه على المتغير الآخر.

معامل التحديد معرف وينتمي إلى المجال التالي:

$$0 \le R^2 \le 1$$

لأن:

$$\sum(\hat{Y}_i - \bar{Y})^2 \le \sum(Y_i - \bar{Y})^2 \Rightarrow \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} \le 1 \Rightarrow R^2 \le 1$$

$$\left.\begin{array}{l}\sum(\hat{Y}_i - \bar{Y})^2 \ge 0 \\ \sum(Y_i - \bar{Y})^2 > 0\end{array}\right\} \Rightarrow R^2 \ge 0$$

1- بالنسبة لنموذج الانحدار الخطي البسيط يكون معامل التحديد هو نفسه مربع معامل الارتباط ما بين متغيرين، أما بالنسبة لنموذج الانحدار المتعدد يصبح هذا التعريف غير صحيح مثلما ما سنرى في الفصل الثاني.

لما يأخذ R^2 أكبر قيمة وهي 1، أي عندما تقع كل نقاط الملاحظات (Y_i, X_i) على الخط المقدر $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$، فالقدرة التفسيرية للنموذج عالية جدا، أي هناك جودة في التوفيق و الارتباط بين المتغير التابع و المستقل.

أما إذا كان R^2 يأخذ أصغر (أسوء) قيمة له وهي الصفر، فليس هناك جودة في التوفيق و الارتباط بين المتغير التابع و المستقل أي ليس للنموذج قدرة تفسيرية على الإطلاق ويعود دلك إلى سببين، إما العلاقة الموجودة بين المتغيرين هي غير خطية أو غياب السببية بينهما.

نذكر أن الفرق الجوهري بين معامل التحديد و معامل الارتباط يكمن في السببية حيث يقيس معامل الارتباط العلاقة بين متغيرين بغض النظر عن الدور الذي يلعبه كل متغير، أما معامل التحديد فيقيس أيضا الارتباط ولكن يأخذ بعين الاعتبار السببية حيث أن المتغير X_i هو الذي يشرح الظاهرة Y_i.

هناك علاقة بين R^2 و $\hat{\beta}_1$، نضع:

$$R^2 = r^2 = \frac{(\text{cov}(X_i, Y_i))^2}{(\sigma_{X_i}\sigma_{Y_i})^2} = \frac{\left(\frac{1}{n}\sum_i (X_i - \bar{X})(Y_i - \bar{Y})\right)^2}{\left(\sqrt{\frac{1}{n}\sum_i (X_i - \bar{X})^2}\sqrt{\frac{1}{n}\sum_i (Y_i - \bar{Y})^2}\right)^2}$$

فنحصل على:

$$R^2 = \frac{\frac{1}{n^2}\left[\sum_i (X_i - \bar{X})(Y_i - \bar{Y})\right]\left[\sum_i (X_i - \bar{X})(Y_i - \bar{Y})\right]}{\frac{1}{n^2}\sum_i (X_i - \bar{X})^2 \sum_i (Y_i - \bar{Y})^2}$$

وعليه:

$$R^2 = \frac{\hat{\beta}_1 \sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (Y_i - \bar{Y})^2}$$

لدينا أيضا:

$$R^2 = \frac{\hat{\beta}_1 \sum (X_i - \bar{X})(Y_i - \bar{Y}) \sum (X_i - \bar{X})^2}{\sum (Y_i - \bar{Y})^2 \sum (X_i - \bar{X})^2} = \frac{\hat{\beta}_1^2 \sum (X_i - \bar{X})^2}{\sum (Y_i - \bar{Y})^2}$$

إذن:

$$R^2=\frac{ESS}{TSS}=\frac{\hat{\beta}_1\sum(X_i-\overline{X})(Y_i-\overline{Y})}{\sum(Y_i-\overline{Y})^2}=\frac{\hat{\beta}_1^2\sum(X_i-\overline{X})^2}{\sum(Y_i-\overline{Y})^2}$$

مثال 3:

باستعمال معطيات المثال الأول، نقوم بتحليل التباين ثم حساب معامل التحديد. نتساءل ما إذا كان لنموذج الاستهلاك قدرة تفسيرية عالية أم لا.

بمعرفة قيمتي TSS و RSS، يمكن حساب قيمة ESS عن طريق معادلة تحليل التباين، حيث:

$$TSS=\sum_i(Y_i-\overline{Y})^2=95.05$$

$$RSS=\sum_{i=1}^{7}\hat{\varepsilon}_i^2=1.32$$

لدينا إذن:

$$ESS=TSS-RSS=95.05-1.32=93.73$$

من خلال الجدول (4)، يمكن تحليل تباين الاستهلاك كما هو مبين في الجدول التالي:

الجدول (4): تحليل التباين باستعمال الانحدار الخطي البسيط

مصدر التغير	مجموع المربعات	درجة الحرية	المربعات المتوسطة
المتغير المستقل	$ESS=93.73$	1	$ESS/1=93.73$
البواقي	$RSS=1.32$	$n-2=7-2=5$	$RSS/5=0.26$
المجموع	$TSS=95.05$	$n-1=7-1=6$	

نقوم بحساب معامل التحديد : R^2

$$R^2=1-\frac{\sum\hat{\varepsilon}_i^2}{\sum(Y_i-\overline{Y})^2}=1-\frac{RSS}{TSS}=1-\frac{1.32}{95.05}=0.9861$$

من خلال نتيجة معامل التحديد، نلاحظ أن الدخل المتاح يفسر الاستهلاك بنسبة 98.61% وبالتالي لنموذج الاستهلاك قدرة تفسيرية عالية.

6. اختبار الفرضيات

بمعرفة توزيع $\hat{\beta}_0$ و $\hat{\beta}_1$ يمكن إجراء اختبار الفرضيات الموضوعة حول معالم النموذج β_0 و β_1 على التوالي. الاختبار الشائع جدا هو فرضية العدم H_0، وتقترح على العموم بأنه لا يوجد أثر على النموذج من قبل متغير مستقل ما، ونظرا إلى أن الباحثين يتمنون قبول النموذج، فإن فرضية العدم توضع عادة لإثبات رفضها إذا أمكن ذلك. ونأمل رفض H_0 بإيجاد القيمة التقديرية والتي تكون تختلف عن الصفر، حتى نقبل النموذج.

1.6. اختبار المعنوية الإحصائية للمعالم

قد يكون النموذج المبني من طرفنا صحيحا أو غير صحيح، و تثبت صحته من خلال اختباره، ويتم ذلك بواسطة فرض معلمة من معالم النموذج تساوي الصفر أو أي عدد آخر، وتسمى فرضية العدم H_0، وما دامت العلاقة بين Y و X قائمة على أساس النموذج الخطي، فإن انعدام هذه العلاقة يعني بأن خط انحدار المجتمع هو عبارة عن خط أفقي، أي $(H_0 : \beta_1 = 0)$ وبما أن الافتراض H_0 خاضع للاختبار، فإنه لا يكون بالضرورة صحيحا، الأمر الذي يتطلب منا وضع فرض بديل $H_1 : \beta_1 \neq 0$. وفي حالة معرفة إشارة β_1 مسبقا من النظرية الاقتصادية فإن الافتراض البديل يكون $H_1 : \beta_1 > 0$ (أو $H_1 : \beta_1 < 0$)، وإذا طلب منا اختبار الفرضية:

$$H_0 : \beta_1 = 0 \quad \text{(فرضية العدم)}$$

ضد: $H_1 : \beta_1 \neq 0$ (الفرضية البديلة)

$$t_c = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}$$

نكتب: $t_c = \dfrac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}}$ وهي القيمة المحسوبة.

ما دمنا نختبر فرضية العدم، نكتب: $t_c = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}$، حيث نقبل H_0 بمستوى معنوية (α%) إذا كانت $\left|\frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}\right| \le t_{n-2,\frac{\alpha}{2}}$ ففي هده الحالة، المعلم β_1 ليس له معنوية إحصائية أي يساوي معنويا الصفر حيث $t_{n-2,\frac{\alpha}{2}}$ مأخوذة من جدول التوزيع t (ستودنت) وتسمى بالقيمة المجدولة، ونرفض H_0 بمستوى معنوية (α%) إذا كانت $\left|\frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}\right| > t_{n-2,\frac{\alpha}{2}}$ أي المعلم β_1 له معنوية إحصائية فهو يختلف معنويا عن الصفر. نقوم بنفس الاختبار مع الثابتة β_0.

إضافة إلى ذلك، عندما يكون حجم العينة كبيرا ($n > 30$) فينبغي استعمال التوزيع الطبيعي و يمكن أخذ القيمة الحرجة $z_{\alpha/2}$ و ذلك بحساب المساحة المظلة للتوزيع الطبيعي.

توجد عدة تساؤلات لدى الباحثين في الاقتصاد القياسي في الاختبار الإحصائي الأفضل بين معامل التحديد R^2 والأخطاء المعيارية للمقدرات، فأيهما أفضل؟ قيمة عالية لـ R^2. أم قيمة منخفضة للأخطاء المعيارية للمقدرات؟ على العموم يكون الاختبار سهلا لما نحصل على قيمة عالية لـ R^2 . وقيمة منخفضة للأخطاء المعيارية، لكن تطبيقيا نادرا ما يحدث ذلك، حيث في أغلب الأحيان نحصل على قيمة عالية لـ R^2. وفي نفس الوقت على قيم عالية للأخطاء المعيارية لبعض المقدرات!؟. ويرى في هذا السياق بعض المتخصصين في الميدان أن تُعطى الأهمية أكثر لقيمة R^2 العالية، ومن ثم يقبلون مقدرات المعالم غير مهتمين بعدم جدية المعنوية الإحصائية لبعض هذه المعالم.

2.6. اختبار التوزيع F (اختبار المعنوية الكلية للنموذج)

إن اختبار معنوية (أثر) المتغير المستقل X_i ($H_0 : \beta_1 = 0$) يمكن أن يكون في شكل توزيع Fisher، حيث لدينا التوزيع الطبيعي المعياري:

$$\hat{\beta}_1 \rightsquigarrow N\left(\beta_1, \frac{\sigma_\varepsilon^2}{\sum_i x_i^2}\right) \Rightarrow \frac{\hat{\beta}_1 - \beta_1}{\sigma_\varepsilon / \sqrt{\sum x_i^2}} \rightsquigarrow N(0,1)$$

يمكن استنتاج أن:

$$\frac{(\hat{\beta}_1 - \beta_1)^2}{\sigma_\varepsilon^2 / \sum x_i^2} \rightsquigarrow \chi^2(1)$$

وما دام $\left(\frac{RSS}{\sigma_\varepsilon^2}\right) \sim \chi^2(1)$، ومستقل توزيعيا عن $\hat{\beta}_1$، فإنه بناءا على تعريف التوزيع F نجد:

$$\frac{\chi^2(1)/1}{\chi^2{}_{(n-2)}/(n-2)} = \frac{(\hat{\beta}_1 - \beta_1)^2 \sum x_i^2}{\sum \hat{\varepsilon}_i^2 /(n-2)} = \frac{(\hat{\beta}_1 - \beta_1)^2 \sum x_i^2}{\hat{\sigma}_\varepsilon^2} \sim F_{1,n-2}$$

وإذا كانت الفرضية $H_0 : \beta_1 = 0$ صحيحة ينتج أن:

$$F = \frac{\hat{\beta}_1^2 \sum x_i^2}{\sum \hat{\varepsilon}_i^2 /(n-2)} = \frac{(n-2)\hat{\beta}_1^2 \sum x_i^2}{RSS} \sim F_{1,n-2}$$

واعتمادا على النتائج السابقة يمكن كتابة الصيغة السابقة من الشكل:

$$F = \frac{\hat{\beta}_1^2 \sum x_i^2}{RSS/(n-2)} = \frac{ESS/1}{RSS/(n-2)} \sim F_{1,n-2}$$

ونقول أننا نرفض $H_0 : \beta_1 = 0$ بمستوى معنوية $(\alpha\%)$ إذا:

$$F_{1,n-2} = \frac{\hat{\beta}_1^2 \sum x_i^2}{RSS/(n-2)} = \frac{ESS/1}{RSS/(n-2)} > F_{\alpha,(1,n-2)}$$

حيث أن $F_{\alpha,(1,n-2)}$ هي القيمة المجدولة، وتؤخذ من جداول توزيع F، وتقبل الفرضية H_0 إذا حدث العكس أي:

$$F_{1,n-2} = \frac{\hat{\beta}_1^2 \sum x_i^2}{RSS/(n-2)} = \frac{ESS/1}{RSS/(n-2)} \leq F_{\alpha,(1,n-2)}$$

وبالمقارنة مع التوزيع t نجد العلاقة التالية:

$$\frac{\hat{\beta}_1 \sqrt{\sum x_i^2}}{\sqrt{RSS/(n-2)}} = \left(\frac{\hat{\beta}_1}{\hat{\sigma}_\varepsilon / \sqrt{\sum x_i^2}}\right)^2 \sim [t_{n-2}]^2 \sim F_{\alpha,(1,n-2)}$$

<u>ملاحظة:</u> تصلح هذه النتيجة لما نختبر المعالم الفردية لنموذج الانحدار فقط.

ولإيجاد العلاق ـة الخاص ـة ب ـالتوزيعين F، t معام ـل التحدي ـد R^2 نع ـود للعلاق ـة:

$$R^2 = r^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

ومنه نكتب:

$$ESS = R^2.TSS = R^2.\sum y_i^2$$
$$RSS = (1-R^2).TSS = (1-R^2).\sum y_i^2$$

ولنعوض ذلك في العلاقة: $F = \frac{ESS/1}{RSS/(n-2)} \sim F_{1,n-2}$ فنجد:

$$F = \frac{R^2/1}{\left(1-R^2\right)/(n-2)} = \frac{R^2}{\left(1-R^2\right)}.(n-2) \sim F_{1,n-2}$$

ونظرا للعلاقة الموجودة ما بين التوزيعين F و t يمكن كتابة: $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2}$.

في توزيع F، نختبر انعدام كل المعالم في آن واحد ضد فرضية معنوية الميل. القيمة المجدولة لإحصائية Fisher في هده الحالة تعتمد على درجتي حرية 1 (في البس ط) و $n-2$ (في المقام).

مثال 4:

باستعمال معطيات المثال الأول، نقوم باختبار المعنوية الإحصائية للمعالم و المعنوية الكلية للنموذج. ندرس أولا إمكانية قبول المقدرات كأس اس للوص ول إلى مع الم المجتم ع الإحصائي، فمثلا هل يختلف الميل الحدي للاستهلاك معنويا عن الصفر؟

لدينا الفرضيتان:

$$H_0 : \beta_1 = 0$$

ضد:

$$H_1 : \beta_1 \neq 0$$

نعلم أن $\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}}$ تتبع توزيع ستيودنت بدرجة حرية $n-2$ و في ظل قبول الفرضية H_0 :

$t_c = \frac{\hat{\beta}_1 - 0}{\hat{\sigma}_{\hat{\beta}_1}} = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}$ تتبع توزيع ستيودنت بدرجة حرية $n-2$.

لدينا إذن: $t_c = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}} = \frac{0.21}{0.011} = 18.81$

نتخذ القرار وذلك بمقارنة القيمة المطلقة للقيمة المحسوبة بالقيمة الحرجة لتوزيع ستيودنت بدرجة حرية 5 و نسبة معنوية 5%. نلاحظ أن $|t_c| = 18.81 > t_{0.025} = 2.57$ و عليه نقبل الفرضية البديلة H_1 أي أن الميل الحدي للاستهلاك يختلف معنويا عن الصفر بنسبة معنوية $\alpha = 5\%$ وهذا يعني أنه يمكن قبول مقدر الميل الحدي للاستهلاك كأساس للوصـ ـول إلى معلمه النظري في المجتمع الإحصائي.

نختبر الآن المعنوية الكلية لنموذج الاستهلاك معتمدا على إحصائية فيشر. لدينا الفرضيتان:

$$H_0 : \beta_0 = \beta_1 = 0$$

ضد: $$H_1 : \beta_1 \neq 0$$

نعلم أيضا أن $\frac{R^2/1}{\left(1-R^2\right)/(n-2)}$ تتبع توزيع فيشر بدرجتي حرية 1 و $n-2$. لـ ـدينا في هذه الحالة:

$$F = \frac{R^2/1}{\left(1-R^2\right)/(n-2)} = \frac{0.9861/1}{(1-0.9861)/(7-2)} = 354.02$$

يتم اتخاذ القرار بمقارنة القيمة المحسوبة بالقيمة الحرجة لتوزيع فيشر بدرجتي حرية 1و 8 و نسبة معنوية 5%. نلاحظ أن $F = 354.02 > F_{0.05}(1,5) = 6.60$ و عليه نقبل الفرضـ ـية البديلة H_1 أي أن للنموذج معنوية إحصائية بنسبة معنوية $\alpha = 5\%$.

7. التنبؤ Forecast:

عقب تقييم نموذج الانحدار والتأكد من استيفاءه للفرضيات والمعايير الإحصائية، يصبح بالإمكان استخدامه لأغراض التنبؤ، وذلك بإيجاد قيم المتغير التابع Y بتغيير قـ ـيم المـ ـتغير المستقل X.

لنأخذ نموذجنا البسيط، ولنفرض أننا نعرف القيمة المستقبلية لـ X. في فترة التنبؤ و نرمز لها بالرمز X_{T+h} ، فإذا فرضنا أن البناء الهيكلي للمعادلة لا يتغير في المستقبل، تكون قيمة المتغير التابع Y في هذه الفترة $T+h$ كما يلي:

$$Y_{T+h} = \beta_0 + \beta_1 X_{T+h} + \varepsilon_{T+h}$$

حيث h يسمى بأفق التنبؤ و Y_{T+h} يعبر عن التنبؤ النظري و T حجم العينة ($t = 1,2,...,T$).

عندما نستعمل علاقة ما للتنبؤ بالقيمة Y ، هناك مصدران لعدم الوضوح والدقة في تنبؤاتنا:

- ❖ عدم معرفتنا للمعلمين β_0 ، β_1 ، وبالتالي يجب الاعتماد على مقدرتي العينة $\hat{\beta}_0$ و $\hat{\beta}_1$ لكي تقدر القيمة Y_{T+h}، إن هذه القيمة هي وسط Y الموافق لـ Y_{T+h} ، أي:

$$E(Y_t) = \beta_0 + \beta_1 X_t$$

$$E(Y_{T+h} \mid X_{T+h}) = \beta_0 + \beta_1 X_{T+h}$$

- ❖ بالإضافة إلى أن الخطأ ε_{t+h} هو متغير عشوائي غير مشاهد، ولهذا حتى وإن عرفنا قيمتي β_0 ، β_1 وبالتالي استطعنا حساب $E(Y_{T+h} \mid X_{T+h})$ ، ونستعين به في التنبؤ، ثم نضع مجال الثقة للتنبؤ لـ Y_{T+h} ، وما دام $E(Y_{T+h} \mid X_{T+h}) = \beta_0 + \beta_1 X_{T+h}$ فيكون المقدر الطبيعي للتنبؤ على الشكل:

$$\hat{Y}_T(h) = \hat{\beta}_0 + \hat{\beta}_1 X_{T+h}$$

ويمكن أن نبين بأن هذا المقدر و الذي يسمى بالتنبؤ التقديري هو مقدر غير متحيز لـ $E(Y_{T+h} \mid X_{T+h})$ ، وأنه من بين المقدرات الخطية غير المتحيزة يعتبر هذا الأخير أحسنها (أي له أصغر تباين)، ويعرف باسم أفضل تنبؤ خطي غير متحيز أي *Best* (*BLUP*) *Linear Unbiased Predictor* ، وإذا فرضنا أن X مستقلة يكون تباين $\hat{Y}_T(h)$ على الشكل:

$$\text{var}\left(\hat{Y}_T(h)\right) = \text{var}\left(\hat{\beta}_0 + \hat{\beta}_1 X_{T+h}\right) = \text{var}\left(\hat{\beta}_0\right) + X_{T+h}^2 . \text{var}\left(\hat{\beta}_1\right) + 2X_{T+h} \text{cov}\left(\hat{\beta}_0, \hat{\beta}_1\right)$$

بما أن:
$$\text{cov}\left(\hat{\beta}_0,\hat{\beta}_1\right)=\frac{-\overline{X}\sigma_{\varepsilon}^2}{\sum x_t^2}$$

فإننا نجد:
$$\text{var}\left(\hat{Y}_T(h)\right)=\sigma_{\varepsilon}^2\left[\frac{1}{T}+\frac{\left(X_{T+h}-\overline{X}\right)^2}{\sum_i\left(X_i-\overline{X}\right)^2}\right]$$

ونلاحظ أن تباين مقدر التنبؤ ينخفض كلما انخفضت القيمة $\left(X_{T+h}-\overline{X}\right)^2$، أي كلما اقتربت X_{T+h} من وسط العينة $\overline{X}$ و ازداد حجم العينة T.

لنأخذ الآن هذا المقدار $E(Y_{T+h}\,|\,X_{T+h})$ كقيمة متنبأ بها (*Predicted Value*) بواسطة تقدير وسطها، إن مقدر الخطأ الداخل في هذا التنبؤ معطى بالعبارة:
$$\hat{\varepsilon}_{T+h}=Y_{T+h}-\hat{Y}_T(h)$$

ونسميه بمقدر خطأ التنبؤ (*Predicted error*) أو (*Forecast error*) ثم نلاحظ أن:
$$E(\hat{\varepsilon}_{T+h})=E(Y_{T+h}-\hat{Y}_T(h))=0$$

ويصبح تباين خطأ التنبؤ:
$$\text{var}(\hat{\varepsilon}_{T+h})=\text{var}(Y_{T+h}-\hat{Y}_T(h))=\text{var}\left(Y_{T+h}\right)+\text{var}\left(\hat{Y}_T(h)\right)$$

إن قيمة Y_{T+h} تعتمد مباشرة على ε_{T+h}، بينما تعتمد $\hat{Y}_T(h)$ على أخطاء العينة $\left(\varepsilon_1,\varepsilon_2,......\varepsilon_T\right)$ بواسطة المقدرين $\hat{\beta}_0$، $\hat{\beta}_1$، وبالتالي يكون $\text{cov}\left(Y_{T+h},\hat{Y}_T(h)\right)=0$.

ونلاحظ أنه إذا كانت X مستقلة فإن: $\text{var}\left(Y_{T+h}\right)=\text{var}(\varepsilon_{T+h})=\sigma_{\varepsilon_{T+h}}^2$ كما وجدنا من قبل.

بما أن:
$$\text{var}\left(\hat{Y}_T(h)\right)=\sigma_{\varepsilon}^2\left[\frac{1}{T}+\frac{\left(X_{T+h}-\overline{X}\right)^2}{\sum_t\left(X_t-\overline{X}\right)^2}\right]$$

فإننا نجد:
$$\text{var}\left(\hat{\varepsilon}_{T+h}\right)=\sigma_{\varepsilon}^2\left[1+\frac{1}{T}+\frac{\left(X_{T+h}-\overline{X}\right)^2}{\sum_t\left(X_t-\overline{X}\right)^2}\right]$$

ولنعرف: $\sigma^2_{\varepsilon_{T+h}} = \mathrm{var}(\hat{\varepsilon}_{T+h})$ لنجد أن:

$$\sigma^2_{\varepsilon_{T+h}} = \sigma^2_{\varepsilon}\left[1+\frac{1}{T}+\frac{\left(X_{T+h}-\overline{X}\right)^2}{\sum_t\left(X_t-\overline{X}\right)^2}\right]$$

ومنه المقدر غير المتحيز لتباين خطأ التنبؤ $\mathrm{var}(\hat{\varepsilon}_{T+h})$ هو:

$$\hat{\sigma}^2_{\varepsilon_{T+h}} = \hat{\sigma}^2_{\varepsilon}\left[1+\frac{1}{T}+\frac{\left(X_{T+h}-\overline{X}\right)^2}{\sum_t\left(X_t-\overline{X}\right)^2}\right]$$

ونلاحظ أنه كلما كبر حجم العينة، $\hat{\sigma}^2_{\varepsilon_{T+h}}$ تقترب من $\hat{\sigma}^2_{\varepsilon}$، أي: $\lim_{n\to\infty}\hat{\sigma}^2_{\varepsilon_{T+h}} = \hat{\sigma}^2_{\varepsilon}$

ولهذا فعندما يكون حجم العينة كبيرا يمكن استعمال $\hat{\sigma}_{\varepsilon}$ كتقريب لـ $\hat{\sigma}_{\varepsilon_{T+h}}$.

نأمل الآن في إيجاد مقياس لتحديد دقة هذا التنبؤ لـ Y_{T+h}، وللقيام بذلك نفترض توزيعا احتماليا معينا للاضطرابات العشوائية، وهو التوزيع الطبيعي. ثم ما دام ε_{T+h} موزعا توزيعا طبيعيا وكذلك Y_{T+h}، كما أن أخطاء العينة $\left(\varepsilon_1,\varepsilon_2,......\varepsilon_T\right)$ موزعة توزيعا طبيعيا، وكذلك $\hat{\beta}_0$، $\hat{\beta}_1$ فإن $\hat{Y}_T(h)$ تكون موزعة طبيعيا أيضا، ولهذا فإن خطأ التنبؤ $\hat{\varepsilon}_{T+h} = Y_{T+h}-\hat{Y}_t(h)$ يكون متغيرا عشوائيا موزعا توزيعا طبيعيا بوسط مساو للصفر وتباين هو $\sigma^2_{\varepsilon_{T+h}}$ أما مقدر هذا التباين فهو $\hat{\sigma}^2_{\varepsilon_{T+h}}$، ومنه:

$$Z = \frac{\hat{\varepsilon}_{T+h}}{\sigma_{\varepsilon_{T+h}}} \sim N(0,1)$$

كما أن $\sigma^2_{\varepsilon_{T+h}}$ تعتمد على القيمة غير المعروفة σ^2_{ε}، فعمليا نعوض بمقدرها $\hat{\sigma}^2_{\varepsilon_{T+h}}$ لتعطي المتغير العشوائي للتوزيع الطبيعي:

$$\frac{\hat{\varepsilon}_{T+h}}{\hat{\sigma}_{\varepsilon_{T+h}}} = \frac{Y_{T+h}-\hat{Y}_T(h)}{\hat{\sigma}_{\varepsilon_{T+h}}} \sim N(0,1)$$

وإذا كانت $z_{\frac{\alpha}{2}}$ هي القيمة الحرجة للتوزيع الطبيعي بحيث تحقق:

$$\Pr\left[-z_{\frac{\alpha}{2}} \le \frac{Y_{T+h}-\hat{Y}_T(h)}{\hat{\sigma}_{\varepsilon_f}} \le +z_{\frac{\alpha}{2}}\right] = 1-\alpha$$

فإن مجال الثقة للتنبؤ يكون: $\hat{Y}_T(h) - \hat{\sigma}_{\varepsilon_f}.z_{\frac{\alpha}{2}} \le Y_{t+h} \le \hat{Y}_T(h) + \hat{\sigma}_{\varepsilon_f}.z_{\frac{\alpha}{2}}$

مثال 5:

نستعمل دائما معطيات المثال الأول للتنبؤ بالاستهلاك لسنة 2008 ثم بناء فترات ثقة للتنبؤ بنسبة معنوية 5%.

إذا علمنا أن القيمة المستقبلية للدخل المتاح لسنة 2008 هي 98 فإنه يمكن التنبؤ بالاستهلاك انطلاقا من النموذج المقدر:

$$\hat{Y}_{2008} = \hat{Y}_7(1) = 0.21X_{2008} + 10.89 = 0.21 \times 98 + 10.89 = 31.47$$

لبناء مجال الثقة للتنبؤ، نحسب أولا تباين خطأ التنبؤ $\mathrm{var}(\hat{\varepsilon}_{T+1})$:

$$\hat{\sigma}^2_{\varepsilon_{2008}} = \hat{\sigma}^2_{\varepsilon}\left[1 + \frac{1}{T} + \frac{\left(X_{2008} - \overline{X}\right)^2}{\sum_t \left(X_t - \overline{X}\right)^2}\right] = 0.26 \times \left[1 + \frac{1}{7} + \frac{(98 - 60.57)^2}{2114.25}\right] = 1.80$$

إذن: $Y_{2008} \in [\hat{Y}_7(1) - t_{0.025}\hat{\sigma}_{\hat{\varepsilon}_{2008}} \ , \ \hat{Y}_7(1) + t_{0.025}\hat{\sigma}_{\hat{\varepsilon}_{2008}}]$

أي: $Y_{2008} \in [31.47 - 2.57\sqrt{1.80} \ , \ 31.47 + 2.57\sqrt{1.80}]$

ومنه: $Y_{2008} \in [28.02 \ , \ 34.91]$

الملحق:

هناك العديد من البرمجيات التي ينبغي على المهتم بالاقتصاد القياسي اسـ ـتعمالها لأنهـ ـا تسمح بقدير الانحدارات و معالجة البيانات الإحصائية. سنعطي لمحة حول كيفية تقـ ـدير النماذج باستعمال برمجيتي RATS و Eviews.

- نقوم بإعطاء نتائج تقدير نموذج الاستهلاك باستعمال RATS "Regression Analysis Time Series" وفق البرنامج التالي:

```
cal 2001 1 1
all 2007:1
data(org=obs,unit=input) / y x
18.47  39.25
19.89  41.84
21.26  49.06
23.71  57.31
25.53  69.89
26.95  78.64
29.48  88.00
```

نقوم بتمثيل الزوج (X_i, Y_i) بيانيا على شكل سحاب نقاط (أنظر الشكل رقم (2-1)):

```
scatter(hlabel='x',vlabel='y') 1
#y x  2001:1 2007:1
```

نقدر النموذج وفق التعليمة التالية للحصول على النتائج المبينة أدناه:

```
linreg(noprint) y / resids
#constant x
```

```
Linear Regression - Estimation by Least Squares
Dependent Variable Y
Annual Data From 2001:01 To 2007:01
Usable Observations      7     Degrees of Freedom     5
Centered R**2     0.986073      R Bar **2   0.983288
Uncentered R**2   0.999669      T x R**2       6.998
Mean of Dependent Variable      23.612857143
Std Error of Dependent Variable  3.980982769
Standard Error of Estimate       0.514640116
Sum of Squared Residuals        1.3242722436
Regression F(1,5)                   354.0249
Significance Level of F           0.00000781
Durbin-Watson Statistic             1.868169

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant              10.857300593  0.705280418     15.39430  0.00002099
2.  X                      0.210591985  0.011192444     18.81555  0.00000781
```

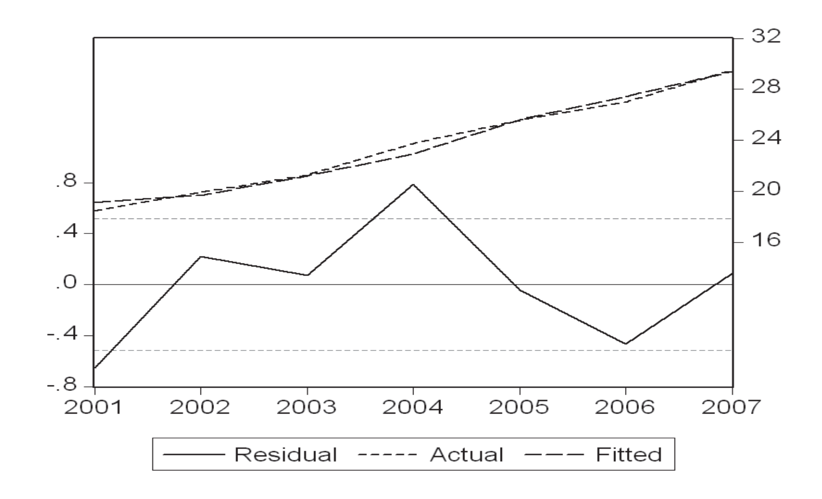
نلاحظ أن القيم المحسوبة ل Student . تعطى مباشرة في RATS كما يعطي أيضا درجة حرية كل من البسط و المقام في إحصائية Fisher. نقبل الفرضية H_1 أي فرضية معنوية المعالم إذا كان الاحتمال P-Value أقل تماما من 0.05، فمثلا نلاحظ أن الاحتمال الخاص بمعامل المتغير المستقل و الذي يساوي إلى 0.00000781 أقل تماما من 0.05 فهذا يعني أن هذا المعلم يختلف معنويا عن الصفر (نقبل H_1).

- أما Eviews فيعطي النتائج مباشرة بطريقة بسيطة:

Dependent Variable: Y
Method: Least Squares
Date: 01/24/09 Time: 15:07
Sample: 2001 2007
Included observations: 7

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	10.85730	0.705280	15.39430	0.0000
X	0.210592	0.011192	18.81555	0.0000

R-squared	0.986073	Mean dependent var	23.61286
Adjusted R-squared	0.983288	S.D. dependent var	3.980983
S.E. of regression	0.514640	Akaike info criterion	1.744259
Sum squared resid	1.324272	Schwarz criterion	1.728804
Log likelihood	-4.104905	F-statistic	354.0249
Durbin-Watson stat	1.868169	Prob(F-statistic)	0.000008



نلاحظ أيضا أن Eviews يعطي مباشرة الشكل البياني الذي يقارن بين الاستهلاك Y_i (Actual) و الاستهلاك المقدر $\hat{Y}_i$ (Fiitted) كما يمثل بيانيا بواقي التقدير $\hat{\varepsilon}_i$ (Residual) الظاهر في أسفل الشكل.

الفَصْلُ الثَّانِي

تحليل الانحدار الخطي العام

الفصل الثاني

تحليل الانحدار الخطي العام

في الواقع الاقتصادي، لا يمكن الاستعانة بالنموذج ذي متغيرين لتحليل الظاهرة الاقتصادية حيث أن هذه الأخيرة لا تفسَّر فقط بمحدد واحد و إنما ينبغي إدماج جميع المحددات أو العوامل المؤثرة في الظاهرة لكي تكون الدراسة أكثر شمولية. في هذا الفصل، نقوم بدراسة الانحدار العام و ذلك باقتراح طريقة لتقدير معالم النموذج و دراسة الخصائص الإحصائية للمقدرات ثم اختبار الفرضيات.

1. الصياغة الرياضية للنموذج الخطي العام:

يستند النموذج الخطي العام على افتراض وجود علاقة خطية ما بين متغير معتمد Y_i وعدد من المتغيرات المستقلة:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_k X_{ik} + \varepsilon_i, \quad i = 1,......,n$$

المتغيرات $X_{i1},..., X_{ij},..., X_{ik}$ تسمى المتغيرات المُفَسِّرة أو المستقلة للمتغير المفسَّر أو التابع Y_i وما يجب ملاحظته أن Y_i مشروح من طرف k متغير مُفَسِّر و لا يمكن لهذه الأخيرة أن تفسر Y بشكل تام، لأنه لا يمكننا في غالب الأحيان حصر جميع الظواهر المؤثرة على Y (بعض الظواهر غير قابلة للتكميم)، لذلك يُدرج حد الخطأ ε_i الذي يتضمن كل المعلومات التي لا تقدمها المتغيرات المفسرة و نفترض عادة بأن المتغيرات المستقلة كلما أخدت بعين الاعتبار كلما كانت المعلومات التي يقدمها الخطأ العشوائي مهملة. نشير فقط إلى أن $\beta_0, \beta_1,...., \beta_k$ هي معالم النموذج، لدينا هنا $(k+1)$ معلم في النموذج.

الـ n . مشاهدة تعطينا n معادلة:

$$i=1: Y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + ... + \beta_k X_{1k} + \varepsilon_1$$
$$i=2: Y_2 = \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + ... + \beta_k X_{2k} + \varepsilon_2$$
....
$$i=n: Y_n = \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + ... + \beta_k X_{nk} + \varepsilon_n$$

يمكن كتابة هذا النظام على الشكل المصفوفي التالي: $Y = X\beta + \varepsilon$

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ \vdots \\ Y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

$Y(n \times 1)$: المتغير التابع أو المفسَّر،

$X(n \times (k+1))$: مصفوفة المتغيرات المُفَسِّرة أو المستقلة،

$\beta((k+1) \times 1)$: شعاع المعالم،

$\varepsilon(n \times 1)$: شعاع الأخطاء.

2. الفرضيات الأساسية للنموذج:

إن بناء نموذج الانحدار الخطي يجب أن يكون مستوفيا لعدد من الفرضيات التي يمكن إجمالها كما يلي:

- الفرضية الأولى: المتغيرات المُفَسِّرة المهملة في النموذج لها أثر متوسط معدوم $E(\varepsilon) = 0$.
- الفرضية الثانية:

$$\begin{cases} \text{var}(\varepsilon_i) = \sigma^2, & \forall i = 1........n \\ \text{cov}(\varepsilon_i, \varepsilon_j) = 0, & \forall i \neq j \end{cases}$$

حيث أن $\text{var}(\varepsilon_i)=\sigma^2 \quad \forall i=1........n$ هي فرضية تجانس التباين "Homoscedasticity" لمختلف الحدود العشوائية، وهذا كفيل بإبعاد الحالة التي تكون فيها الأخطاء تتبع تغيرات قيم المتغيرات المفسرة و $Cov(\varepsilon_i,\varepsilon_j)=0 \quad \forall i\neq j$، أي أن الأخطاء ليست مرتبطة ببعضها، وأن نتيجة تجربة لا تؤثر على بقية النتائج. يمكن كتابة هاتين الفرضيتين على الشكل المصفوفي:

$$\Omega_\varepsilon = E(\varepsilon\varepsilon') = \begin{pmatrix} \sigma_\varepsilon^2 & 0 & \cdots & 0 \\ 0 & \sigma_\varepsilon^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_\varepsilon^2 \end{pmatrix} = \sigma_\varepsilon^2 I_n$$

تسمى المصفوفة Ω_ε مصفوفة التباين– التباين المشترك للأخطاء.

- ❖ الفرضية الثالثة: المصفوفة X غير عشوائية وثابتة: تعني بأن قيم المتغيرات المستقلة يمكن مراقبتها، وبالإضافة إلى ذلك نفترض X ثابتة لضمان بأن قيم المتغيرات المستقلة لا تتغير من حين لآخر، أي ;

$$\text{cov}(X,\varepsilon) = E(X'\varepsilon) = 0$$

- ❖ الفرضية الرابعة: عدد المشاهدات n هو أكبر من عدد المتغيرات المفسرة k ، وهي الحالة التي تلغي الارتباط الخطي بين المتغيرات المستقلة.

3. تقدير شعاع المعالم β وتباين الأخطاء σ^2:

في النموذج $Y = X\beta + \varepsilon$، المجاهيل الوحيدة هي β و ε، المصفوفة X و الشعاع Y هي معطيات النموذج، ويجب الإشارة إلى أن شعاع الأخطاء غير مشاهد ولذلك حتى معرفة قيمة β لا تسمح للمتغيرات المستقلة بإعطاء القيمة الحقيقية لـ Y. بالضبط.

وعلينا إذن تقدير β بشكل يجعل $\hat{Y}$ أقرب ما يمكن للمتغير التابع Y، ولهذا الغرض توجد عدة طرق، فيما نستعرض نحن طريقتي المربعات الصغرى والمعقولية العظمى.

1.3. طريقة المربعات الصغرى:

1.1.3. تقدير شعاع المعالم β :

تهدف هذه الطريقة إلى إيجاد تقدير للشعاع β الذي يُصَغِّر مجموع مربعات الانحراف $\hat{\varepsilon}_i$ بين القيمة المقدرة $\hat{Y}$ والقيمة الحقيقية Y.

$$\Rightarrow \quad \hat{\varepsilon} = Y - \hat{Y} = \begin{pmatrix} \hat{\varepsilon}_1 \\ \hat{\varepsilon}_2 \\ \vdots \\ \vdots \\ \hat{\varepsilon}_n \end{pmatrix}$$

$$Min\sum_{i=1}^{n} \hat{\varepsilon}_i^2 = Min\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$$

$$\hat{\varepsilon}_i = Y_i - \hat{Y}_i \qquad i = 1.......n$$

$$Min\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 = Min\left(Y - \hat{Y}\right)'\left(Y - \hat{Y}\right) = Min\,\hat{\varepsilon}'\hat{\varepsilon}$$

نسمي:

$$\Gamma\left(Y, X, \hat{\beta}\right) = \left(Y - \hat{Y}\right)'\left(Y - \hat{Y}\right) = \hat{Y}'\hat{Y} - 2\hat{Y}'Y + Y'Y = \hat{\beta}'X'X\hat{\beta} - 2\hat{\beta}'X'Y + Y'Y$$

حيث: $\hat{Y} = X\hat{\beta}$. ومنه الهدف هو

$$\min_{\hat{\beta}} \Gamma\left(Y, X, \hat{\beta}\right)$$

وإذا كان $\hat{\beta}$ موجود فيجب أن يحقق الشرط الضروري:

$$\frac{\partial\Gamma\left(Y, X, \hat{\beta}\right)}{\partial\hat{\beta}'} = 0 \Leftrightarrow 2(X'X)\hat{\beta} - 2X'Y = 0$$

وبما أن رتبة X هي $k+1$ فإن: $\left(X'X\right)$ مصفوفة مربعة $\left((k+1)\times(k+1)\right)$ رتبتها $k+1$ وتقبل معكوس $\left(X'X\right)^{-1}$.

ومنه: $\quad 2(X'X)\hat{\beta} - 2X'Y = 0 \quad \Rightarrow (X'X)\hat{\beta} - X'Y = 0$

نضرب طرفي المعادلة بـ $\left(X'X\right)^{-1}$. لنحصل على: $\hat{\beta} = (X'X)^{-1}X'Y$ وهو تقدير لـ β .

وللتأكد من أن $\hat{\beta}$ المتحصل عليه هو قيمة دنيا لـ $\Gamma\left(Y,X,\hat{\beta}\right)$، يجب تحقيق الشرط من الدرجة الثانية:

$$\frac{\partial^2\Gamma\left(Y,X,\hat{\beta}\right)}{\partial\hat{\beta}'\partial\hat{\beta}}=\left(X'X\right)>0$$

وهي مصفوفة معرفة موجبة ومنه فإن $\hat{\beta}$ هو نهاية صغرى.

والآن لنرمز بـ A للمصفوفة $(X'X)^{-1}X'$، حيث:

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{pmatrix} \Rightarrow \hat{\beta}=A.Y \quad \therefore \hat{\beta}_i=\sum_{j=1}^{n}a_{ij}Y_j \quad ,i=1......k$$

ومنه نرى أن مختلف المقدرات $\left(\hat{\beta}_k.......\hat{\beta}_2,\hat{\beta}_1\right)$ هي على شكل خطي مع المتغير Y.

كذلك لدينا: $\hat{\beta}=(X'X)^{-1}X'Y$

و أيضا: $Y=X\beta+\varepsilon$

إذن:

$$\hat{\beta}=(X'X)^{-1}X'[X\beta+\varepsilon]=(X'X)^{-1}X'X\beta+(X'X)^{-1}X'\varepsilon\Rightarrow\hat{\beta}=\beta+(X'X)^{-1}X'\varepsilon$$

بإدخال التوقع الرياضي:

$$E(\hat{\beta})=\beta+(X'X)^{-1}X'E(\varepsilon) \quad / \quad E(\varepsilon)=0$$

نحصل في الأخير: $E(\hat{\beta})=\beta$

نستنتج أن التقدير $\hat{\beta}$ لـ β المحصل عليه بطريقة المربعات الصغرى غير متحيز. بالإضافة إلى ذلك فإن $\hat{\beta}$ هو التقدير الأفضل من ضمن كل التقديرات الخطية غير المتحيزة لـ β $(BLUE)$.

2.1.3. تقدير تباين الأخطاء σ^2 و مصفوفة التباين-التباين المشترك للمقدرات $\Omega_{\hat{\beta}}$:

إحدى فرضيات النموذج هي $E(\varepsilon\varepsilon')=\Omega_\varepsilon=\sigma^2I_n$ وبما أن σ^2 غير معروف، فينبغي تقديره:

$$\hat{\varepsilon} = Y - X\hat{\beta} = X\beta + \varepsilon - X\hat{\beta} = \varepsilon - X(\hat{\beta} - \beta) = \varepsilon - X(X'X)^{-1}X'\varepsilon = (I_n - X(X'X)^{-1}X')\varepsilon$$

نضع: $M_X = (I_n - X(X'X)^{-1}X')$ ، حيث M_X *تسمى المصفوفة الدورية* أي:

$$M_X = M_X^{'} M_X = M_X^2 = M'$$

بالإضافة إلى ذلك:

$$M_X X = 0$$

ومنه: $\hat{\varepsilon}^{'}\hat{\varepsilon} = \varepsilon' M_X^{'} M_X \varepsilon$ أي: $\hat{\varepsilon}^{'}\hat{\varepsilon} = \varepsilon' M_X \varepsilon$

ندخل التوقع الرياضي على الطرفين: $E(\hat{\varepsilon}^{'}\hat{\varepsilon}) = E(\varepsilon' M_X \varepsilon)$

ويجب الملاحظة أن أثر $\hat{\varepsilon}^{'}\hat{\varepsilon}$ يساوي أثر $\varepsilon' M\varepsilon$، ونعلم أيضا أن أثر (AB)=أثر (BA).

يكون لدينا إذن: أثر ($\varepsilon' M\varepsilon$)=أثر ($\varepsilon\varepsilon' M$)

$$E(\hat{\varepsilon}^{'}\hat{\varepsilon}) = E(\varepsilon'\varepsilon) Tr(M_X)$$

نعلم أن: $E(\varepsilon'\varepsilon) = \sigma^2$ وعليه: $E(\hat{\varepsilon}^{'}\hat{\varepsilon}) = \sigma^2 \{Tr(I_n) - Tr(X(X'X)^{-1}X')\}$

ومنه:

$$E(\hat{\varepsilon}^{'}\hat{\varepsilon}) = \sigma^2 (n - k - 1)$$

حيث: $Tr(X(X'X)^{-1}X') = k + 1$ ؛ $Tr(I_n) = n$

لكي نحصل على تقدير غير متحيز لـ σ^2. يكفي قسمة العبارة على $(n-k-1)$:

$$E\left(\frac{\hat{\varepsilon}^{'}\hat{\varepsilon}}{n-k-1}\right) = \sigma^2$$

في حالة الانحدار المتعدد حيث هناك $k+1$ معلم للتقدير و n عدد المشاهدات، وهذا يُعطي عدد درجات الحرية $n-k-1$، إذن:

$$\hat{\sigma}_{\varepsilon}^2 = \frac{\hat{\varepsilon}^{'}\hat{\varepsilon}}{n-k-1} = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n-k-1}$$

لقد برهننا أن $\hat{\beta} - \beta$ تساوي $(X'X)^{-1}X'\varepsilon$ وإحدى فرضيات النموذج هي $E(\varepsilon\varepsilon^{'}) = \Omega_{\varepsilon} = \sigma^2 I_n$. لدينا:

$$\hat{\beta} - \beta = (X'X)^{-1}X'\varepsilon$$

نقوم بحساب $(\hat{\beta} - \beta)(\hat{\beta} - \beta)^{'}$، حيث:

$$(\hat{\beta} - \beta)(\hat{\beta} - \beta)^{'} = (X'X)^{-1}X'\varepsilon\varepsilon^{'}X(X'X)^{-1}$$

بإدخال التوقع الرياضي على الطرفين، نتحصل على مصفوفة التباين المشترك للمقدرات:

$$\Omega_{\hat{\beta}} = E((\hat{\beta}-\beta)(\hat{\beta}-\beta)') = (X'X)^{-1}X'E(\varepsilon\varepsilon')X(X'X)^{-1} = (X'X)^{-1}X'\Omega_{\varepsilon}X(X'X)^{-1}$$

إذن:

$$\Omega_{\hat{\beta}} = \sigma_{\varepsilon}^2(X'X)^{-1}$$

بما أن σ_{ε}^2 غير معروف، فانه يمكن استبداله بمقدر تباين الأخطاء $\hat{\sigma}_{\varepsilon}^2$ وعليه:

$$\hat{\Omega}_{\hat{\beta}} = \hat{\sigma}_{\varepsilon}^2(X'X)^{-1}$$

مثال 1:

لدينا البيانات المبينة في الجدول التالي:

الجدول (1): البيانات الإحصائية لـ Y، X_1 و X_2.

X_{i2}	X_{i1}	Y_i	i
3	2	4	1
7	4	6	2
10	5	7	3
8	7	9	4
8	9	10	5
9	10	12	6
11	12	14	7
13	14	16	8
14	15	18	9
15	17	20	10

لدينا 10 مشاهدات و متغيران مستقلان، يكتب النموذج كما يلي:

$$Y = X\beta + \varepsilon$$

مع:

$$Y = \begin{pmatrix} 4 \\ 6 \\ 7 \\ \vdots \\ 20 \end{pmatrix};\ X = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 7 \\ 1 & 5 & 10 \\ \vdots & \vdots & \vdots \\ 1 & 17 & 15 \end{pmatrix};\ \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix};\ \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_{10} \end{pmatrix}$$

نعلم أن $\hat{\beta}=(X'X)^{-1}X'Y$. نقوم إذن بحساب $X'X$ و $(X'X)^{-1}$ ثم $X'Y$

$$X'X=\begin{pmatrix}1&1&1&\cdots&1\\2&4&5&\cdots&17\\3&7&10&\cdots&15\end{pmatrix}\begin{pmatrix}1&2&3\\1&4&7\\1&5&10\\\vdots&\vdots&\vdots\\1&17&15\end{pmatrix}=\begin{pmatrix}10&95&98\\95&1129&1081\\98&1081&1078\end{pmatrix}$$

و منه:

$$(X'X)^{-1}=\begin{pmatrix}1.1725&0.0852&-0.1921\\0.0852&0.0284&-0.0362\\-0.1921&-0.0362&0.0547\end{pmatrix}$$

و أيضا:

$$X'Y=\begin{pmatrix}1&1&1&\cdots&1\\2&4&5&\cdots&17\\3&7&10&\cdots&15\end{pmatrix}\begin{pmatrix}4\\6\\7\\\vdots\\20\end{pmatrix}=\begin{pmatrix}116\\1342\\1298\end{pmatrix}$$

لدينا:

$$\hat{\beta}=\begin{pmatrix}\hat{\beta}_0\\\hat{\beta}_1\\\hat{\beta}_2\end{pmatrix}=(X'X)^{-1}X'Y=\begin{pmatrix}1.1725&0.0852&-0.1921\\0.0852&0.0284&-0.0362\\-0.1921&-0.0362&0.0547\end{pmatrix}\begin{pmatrix}116\\1342\\1298\end{pmatrix}=\begin{pmatrix}1.0996\\0.9776\\0.1237\end{pmatrix}$$

لإيجاد مصفوفة التباين-التباين المشترك للمقدرات، ينبغي أولا حساب تباين البواقي حيث يتم حساب قيم $\hat{Y}_i$ انطلاقا من الانحدار الخطي:

$\hat{Y}_i=1.0996+0.9776X_{i1}+0.1237X_{i2}$ و بواقي التقدير $\hat{\varepsilon}_i$ من المعادلة $\hat{\varepsilon}_i=Y_i-\hat{Y}_i$.

الجدول (2): حساب البواقي

$\hat{\varepsilon}_i^2$	$\hat{\varepsilon}_i$	$\hat{Y}_i$	Y_i	i
0.3292	0.5738	3.4261	4	1
0.0152	0.1235	5.8764	6	2
0.0507	-0.2252	7.2252	7	3
0.0044	0.0668	8.9331	9	4
0.7893	-0.8884	10.8884	10	5
0.0001	0.0101	11.9898	12	6
0.0371	-0.1926	14.1926	14	7
0.1563	-0.3954	16.3954	16	8
0.2532	0.5031	17.4968	18	9
0.1798	0.4241	19.5758	20	10
1.8157	0			المجموع

لدينا:
$$\hat{\sigma}_\varepsilon^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k-1} = \frac{\sum_{i=1}^{10}\hat{\varepsilon}_i^2}{n-k-1} = \frac{1.8157}{10-2-1} = 0.2593$$

نقوم بتحديد مصفوفة التباين–التباين المشترك للمقدرات $\hat{\Omega}_{\hat{\beta}}$، ليكن:

$$\hat{\Omega}_{\hat{\beta}} = \hat{\sigma}_\varepsilon^2 (X'X)^{-1}$$

$$\hat{\Omega}_{\hat{\beta}} = 0.2593 \times \begin{pmatrix} 1.1725 & 0.0852 & -0.1921 \\ 0.0852 & 0.0284 & -0.0362 \\ -0.1921 & -0.0362 & 0.0547 \end{pmatrix}$$

نجد تباين كل مقدر بضرب تباين البواقي بكل عنصر من عناصر قطر المصفوفة $(X'X)^{-1}$:

$$\hat{\sigma}_{\hat{\beta}_0}^2 = 0.2593 \times 1.1725 = 0.3040 \;\rightarrow\; \hat{\sigma}_{\hat{\beta}_0} = 0.5513$$

$$\hat{\sigma}_{\hat{\beta}_1}^2 = 0.2593 \times 0.0284 = 0.0073 \;\rightarrow\; \hat{\sigma}_{\hat{\beta}_1} = 0.0858$$

$$\hat{\sigma}_{\hat{\beta}_2}^2 = 0.2593 \times 0.0547 = 0.0141 \;\rightarrow\; \hat{\sigma}_{\hat{\beta}_2} = 0.1190$$

2.3. طريقة المعقولية العظمى:

بفرض أن الأخطاء تتوزع توزيعا طبيعيا حيث $\varepsilon_i \sim N(0,\sigma^2), \quad i = 1........n$ فان دالة كثافتها تكتب على الشكل التالي:

$$f(\varepsilon_i)=\frac{1}{(2\pi\sigma^2)^{1/2}}.\exp\left\{-\frac{1}{2}.\frac{\varepsilon_i^2}{\sigma^2}\right\}$$

نسمي $L(\varepsilon_1,...,\varepsilon_n)$ دالة المعقولية، حيث:

$$L(\varepsilon_1,...,\varepsilon_n)=\prod_{i=1}^{n}f(\varepsilon_i)=\prod_{i=1}^{n}(2\pi\sigma^2)^{-1/2}\exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right)$$

ولكن: $\varepsilon_i=Y_i-\beta_0-\beta_1X_{i1}-....-\beta_kX_{ik}$ أي: $\varepsilon=Y-X\beta$

لدينا إذن:

$$L(\beta,\sigma^2)=(2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2}\frac{(Y-X\beta)'(Y-X\beta)}{\sigma^2}\right)$$

لنأخذ لوغاريتم دالة المعقولية:

$$\ln L(\beta,\sigma^2)=-\frac{n}{2}\ln(2\pi\sigma^2)-\frac{1}{2\sigma^2}(Y-X\beta)'(Y-X\beta)$$

يتم تقدير معالم النموذج وتباين الأخطاء ودلك بتعظيم لوغاريتم دالة المعقولية العظمى، حيث:

$$\max_{\beta,\sigma^2}\{\ln L(\beta,\sigma^2)\}=\max_{\beta,\sigma^2}\left\{-\frac{n}{2}\ln(2\pi\sigma^2)-\frac{1}{2\sigma^2}(Y-X\beta)'(Y-X\beta)\right\}$$

الشروط الضرورية لكي يكون لوغاريتم الدالة عند قيمتها العظمى هي:

$$\frac{\partial\ln L}{\partial\beta}=0\Leftrightarrow-\frac{1}{2\sigma^2}\frac{\partial}{\partial\beta}\left(Y'Y+\beta'X'X\beta-2\beta'X'Y\right)\Leftrightarrow\frac{1}{2\sigma^2}\{2X'X\beta-2X'Y\}=0$$

$$\frac{\partial\ln L}{\partial\sigma^2}=0\Leftrightarrow-\frac{n}{2}\frac{1}{\sigma^2}+\frac{1}{2}\frac{1}{\sigma^4}(Y-X\beta)'(Y-X\beta)=0$$

إذا كانت $X'X$ رتبتها $k+1$ يعطينا: $\hat{\beta}_{MLE}=(X'X)^{-1}X'Y$

و $\hat{\sigma}^2_{MLE}=\frac{1}{n}\left(Y-X\hat{\beta}_{MLE}\right)'\left(Y-X\hat{\beta}_{MLE}\right)$

أي:
$$\hat{\sigma}^2_{MLE} = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n}$$

ينبغي القول أن المقدر $\hat{\sigma}^2_{MLE}$ متحيز، حيث لدينا:

$$\hat{\sigma}^2_{MLE} = \left(\frac{n-k-1}{n}\right)\left(\frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k-1}\right) \Rightarrow \hat{\sigma}^2_{MLE} = \left(\frac{n-k-1}{n}\right)\hat{\sigma}^2_{OLS}$$

حيث $\hat{\sigma}^2_{OLS}$ المقدر غير المتحيز لتباين الأخطاء بطريقة المربعات الصغرى، أي: $E(\hat{\sigma}^2_{OLS}) = \sigma^2$، وهذا يؤدي إلى:

$$E(\hat{\sigma}^2_{MLE}) = \left(\frac{n-k-1}{n}\right)\sigma^2 \neq \sigma^2$$

يكون $\hat{\sigma}^2_{MLE}$ أحسن تقدير لـ σ^2. إذا كان حجم العينة كبيرا أي $\frac{n-k-1}{n} \to 1$ لما $n \to \infty$ نستنتج أن: $E(\hat{\sigma}^2_{MLE}) = \sigma^2$.

4. اختبار جودة التوفيق والارتباط:

عندما يكون لدينا أكثر من متغير مستقل في نموذج الانحدار الخطي، ننتقل من معامل التحديد العادي (معامل الارتباط البسيط) إلى معامل التحديد المضاعف، وفي حين أن الأول يقيس العلاقة بين متغير مستقل وآخر تابع، فإن الثاني وبالإضافة إلى نفس الدور فإنه يمكن أن يدرس العلاقة بين المتغير التابع Y وعدة متغيرات مستقلة مرة واحدة، ويسمى بمعامل التحديد المتعدد. كما أنه يمكن أن نبين العلاقة بين متغير مستقل وعدة متغيرات مستقلة أخرى بواسطة معامل يسمى بمعامل الارتباط المتعدد، ويستعمل عادة في اختبارات اكتشاف التعدد الخطي، حيث يعتمد عليه الباحثان Farrar-Glauber في شكل معاملات تحديد جزئية على شكل $R^2_{X_j,X_1,X_2,......X_k}$ حيث أنه يربط ما بين المتغير المستقل X_j وبقية المتغيرات المستقلة الأخرى من غير X_j.

أما معامل التحديد المتعدد R^2 فهو يشير إلى النسبة التي يمكن تفسيرها من التغير الكلي في المتغير التابع Y بدلالة المتغيرات المستقلة المدرجة في المعادلة، ويستعمل كمقياس لجودة التوفيق في نموذج الانحدار المحتوي على k متغير مستقل، ولحسابه يمكن إتباع نفس الطريقة المستعملة في النموذج الخطي البسيط: $TSS=ESS+RSS$ ففي النموذج ذي k متغير مستقل:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + ... + \beta_k X_{ik} + \varepsilon_i \quad i = 1,....,n$$

يمكن حساب R^2 على الشكل:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum \hat{\varepsilon}_i^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2} = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}$$

أما إذا كان النموذج لا يحتوي على ثابتة، فإن R^2 يكتب بدون تركيز المتغيرات:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\hat{\varepsilon}'\hat{\varepsilon}}{Y'Y} = \frac{\hat{Y}'\hat{Y}}{Y'Y}$$

وتتراوح قيمة R^2 بين 0 (عندما لا تُفَسِّر معادلة الانحدار أيا من التغير في Y)، و1 (عندما تقع كل النقاط على خط الانحدار).

هناك علاقة بين معامل التحديد و شعاع المقدرات:

$$R^2 = \frac{\hat{Y}'\hat{Y}}{Y'Y} = \frac{\hat{\beta}'X'X\hat{\beta}}{Y'Y}$$

معامل التحديد يؤول أيضا إلى العلاقة التالية:

$$R^2 = \frac{\hat{\beta}_1 \sum x_{i1} y_i + \hat{\beta}_2 \sum x_{i2} y_i + ... + \hat{\beta}_k \sum x_{ik} y_i}{\sum y_i^2}$$

حيث: $y_i = Y_i - \overline{Y} \quad , \quad x_{ij} = X_{ij} - \overline{X}_j \qquad \forall j = 1......k \, , \quad \forall i = 1......n$

إذا كان النموذج لا يحتوي على ثابتة، فإننا نعوض شعاع المقدرات بما يساويه، أي:

$\hat{\beta} = (X'X)^{-1} X'Y$ ، إذن: $$R^2 = \frac{\hat{\beta}' X' X\hat{\beta}}{Y'Y} = \frac{\hat{\beta}' X' Y}{Y'Y}$$

ولكن هناك مجموعة من المشاكل نواجهها مع استعمال R^2 منها:

- ❖ أولا: كل نتائجنا الإحصائية تأتي من الفرضية القائلة بأن نموذجنا المبني في المعادلة $Y = X\beta + \varepsilon$ يكون صحيحا، ثم ليس لدينا طريقة أو قيمة إحصائية بديلة للمقارنة.
- ❖ ثانيا: إن R^2 غير حساس لعدد المتغيرات المستقلة والموجودة بالنموذج، حيث إن إضافة متغيرات مستقلة أخرى لمعادلة الانحدار لا يمكن أبدا أن تُقلل من قيمة R^2، وبالعكس فإنها يمكن أن تزيد من قيمته (لأن إضافة متغير مستقل جديد للنموذج لا يؤثر في التغيرات الكلية TSS، بينما يزيد في قيمة الانحرافات المشروحة ESS)، ويصبح تفسير واستعمال R^2 صعبا عندما يكون النموذج بدون الحد الثابت، حيث ليس بالضرورة في هذه الحالة أن يكون محصورا بين 0 و1.

إن الصعوبات في استعمال R^2 كمقياس لجودة التوفيق راجعة لأن هذا المعامل يعتمد على التغيرات الحاصلة في Y (المشروحة وغير المشروحة)، وبالتالي فإنه لا يأخذ بعين الاعتبار عدد درجات الحرية في أي مشكل إحصائي. ولهذا الغرض يُستعمل معامل أخر يسمى معامل التحديد المصحح $\overline{R}^2$.

فإذا كان تعريف R^2 هو: $R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$ ، فإن تعريف $\overline{R}^2$ هو:

$$\overline{R}^2 = 1 - \frac{RSS/(n-k-1)}{TSS/(n-1)}$$

حيث n: عدد المشاهدات و $k+1$: عدد المعالم المقدرة. وبتعويض بسيط نجد:

$$\overline{R}^2 = 1 - (1 - R^2)\left(\frac{n-k-1}{n-1}\right)$$

ومن المعادلة الأخيرة أعلاه، تظهر العلاقة بين R^2 و $\overline{R}^2$ حيث أن:

1. $R^2 \geq \overline{R}^2$ إذا كانت $k > 1$
2. $R^2 = \overline{R}^2$ إذا كانت $k = 1$
3. $\overline{R}^2$ يمكن أن يأخذ قيما سالبة.

إذا كان حجم العينة n كبيرا، فإن R^2 و $\overline{R}^2$ يقتربان قي قيمتهما، لكن في العينات الصغيرة، إذا كان عدد المتغيرات المستقلة كبيرا بالمقارنة مع حجم العينة، فإن $\overline{R}^2$ يقل بكثير على R^2، ويمكن أن يأخذ قيما سالبة، في هذه الحالة يجب شرحه على أساس أن قيمته تساوي الصفر.

إذن $\overline{R}^2$ له مجموعة من الخصائص تجعله وسيلة قياس جودة التوفيق أفضل من R^2، فهو على الأقل يُجيب على تساؤلات بعض الباحثين حول أهمية زيادة عدد المتغيرات للنموذج، بدون التفكير في سبب ظهور هذه المتغيرات على كل حال، رغم ذلك لا يجب التفكير في أن $\overline{R}^2$ يحل كل المشاكل المتعلقة بالمقياس R^2 لجودة التوفيق، حيث أن القرار حول إمكانية ظهور بعض المتغيرات في النموذج أم لا، تبقى معتمدة على اعتبارات نظرية أخرى في القياس الاقتصادي، كما أن القيمة العددية لـ $\overline{R}^2$. تكون جد حساسة لنوع المعطيات أو البيانات المستعملة.

مثال 2:

نفس معطيات المثال السابق، المطلوب حساب معامل التحديد R^2 و معامل التحديد المصحح $\overline{R}^2$. نعلم أن $\sum_{i=1}^{10} \hat{\varepsilon}_i^2 = 1.8157$، فينبغي حساب $\sum_{i=1}^{10} \left(Y_i - \overline{Y}\right)^2 = 256.4$

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^{10} \hat{\varepsilon}_i^2}{\sum_{i=1}^{10} \left(Y_i - \overline{Y}\right)^2} = 1 - \frac{1.8157}{256.4} = 0.9929$$

و

$$\overline{R}^2 = 1-\left(1-R^2\right)\left(\frac{n-k-1}{n-1}\right) = 1-(1-0.9929)\left(\frac{10-2-1}{10-1}\right) = 0.9908$$

من الملاحظ أن معامل التحديد R^2 يبقى أقل نسبيا من معامل التحديد المصحح $\overline{R}^2$.

من خلال نتائج معامل التحديد، نستنتج أن للنموذج قدرة تفسيرية عالية جدا أي أن المتغيرات المستقلة تشرح المتغير التابع بنسبة 99.28%.

5. اختبار الفرضيات

1.5. اختبار المعنوية الإحصائية للمعالم

بإدخال قانون التوزيع الطبيعي المتعدد ونظرا إلى أن $\hat{\beta}$ هو دالة خطية لشعاع الأخطاء العشوائية، فإن هذا المتغير له صفة المتغير العشوائي ويتبع كذلك قانون التوزيع الطبيعي المتعدد

نضع: $A = (X'X)^{-1}X'$

لدينا: $\hat{\beta} = \beta + A\varepsilon$

ومنه فإن: $\hat{\beta} \sim N\left(\beta, \sigma_\varepsilon^2 (X'X)^{-1}\right)$

ثم لدينا بواقي المربعات الصغرى: $\hat{\varepsilon} = M_X \varepsilon$

إذ أن: $\hat{\varepsilon}'\hat{\varepsilon} = \varepsilon' M_X \varepsilon$

مع: $M_X = (I - X(X'X)^{-1}X')$

ومنه: $\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma_\varepsilon^2} = \frac{\varepsilon' M_X \varepsilon}{\sigma_\varepsilon^2} = \frac{(n-k-1)\hat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} \sim \chi^2_{n-k-1}$

مع الخاصية $M_X X = 0$ يكون الشعاعان $\hat{\beta}$ و $\hat{\varepsilon}$ يتبعان التوزيع الطبيعي المتعدد ومستقلين عن بعضهما البعض، وبالتالي فهما شعاعان متعامدان حيث:

$$\text{cov}\left(\hat{\varepsilon}, \hat{\beta}\right) = E\left[\hat{\varepsilon}\left(\hat{\beta}-\beta\right)'\right] = E\left[M_X \varepsilon\varepsilon' A'\right] = \sigma_\varepsilon^2 M_X A = 0, \quad M_X X = 0$$

ومنه نستنتج أن شعاع المقدرات $\hat{\beta}$ مستقل كذلك عن $\hat{\varepsilon}'\hat{\varepsilon}$، والذي يستلزم أن $\hat{\beta}$ موزع استقلاليا عن $\frac{RSS}{\sigma_\varepsilon^2}$ أو $\hat{\sigma}_\varepsilon^2$، ونكتب:

$$\hat{\beta}_j \sim N\left(\beta_j, \sigma_\varepsilon^2 a_{jj}\right), \quad j = 0,1,....,k$$

حيث أن a_{jj} هو العنصر j الموجود بقطر المصفوفة AA' (أو $(X'X)^{-1}$) ، مع $A = (X'X)^{-1}X'$.

ولدينا كذلك: $\left(\hat{\beta}_j - \beta_j\right) \sim N\left(0, \sigma_\varepsilon^2 a_{jj}\right), \quad j = 0,1,....,k$

ومنه: $\left(\frac{\hat{\beta}_j - \beta_j}{\sigma_\varepsilon \sqrt{a_{jj}}}\right) \sim N(0,1), \quad j = 0,1,....,k$

وليصبح قانون التوزيع t على الشكل:

$$t = \frac{N(0,1)}{\sqrt{\chi_{n-k}^2 / (n-k-1)}} = \frac{\dfrac{\hat{\beta}_j - \beta_j}{\sigma_\varepsilon \sqrt{a_{jj}}}}{\sqrt{\dfrac{(n-k-1)\hat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} \Big/ (n-k-1)}}$$

ونجد بعد الاختصار: $t = \frac{\hat{\beta}_j - \beta_j}{\sigma^2 \sqrt{a_{jj}}} = \frac{\hat{\beta}_j - \beta_j}{\sigma_{\hat{\beta}_j}} \sim t_{n-k-1}$

تساعدنا هذه المعادلة إذن على تكوين مجالات الثقة لمعالم النموذج بنفس الطريقة المذكورة في حالة النموذج البسيط،

$H_0 : \beta_j = 0$ (فرضية العدم)

$j = 0,1,...,k$ ضد : $H_1 : \beta_j \neq 0$ (الفرضية البديلة)

نكتب: $t_c = \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_{\hat{\beta}_j}}$ وهي القيمة المحسوبة.

ما دمنا نختبر فرضية العدم، نكتب $t_c = \frac{\hat{\beta}_j}{\hat{\sigma}_{\hat{\beta}_j}}$، حيث نقبل H_0 بمستوى معنوية α إذا كانت $\left|\frac{\hat{\beta}_j}{\hat{\sigma}_{\hat{\beta}_j}}\right| \leq t_{n-k-1,\frac{\alpha}{2}}$ ففي هده الحالة، المعلم β_j ليس له معنوية إحصائية أي يساوي

معنويا الصفر حيث $t_{n-k-1,\frac{\alpha}{2}}$ مأخوذة من جدول التوزيع t، ونرفض H_0 بمستوى معنوية α

إذا كانت $\left|\frac{\hat{\beta}_j}{\hat{\sigma}_{\hat{\beta}_j}}\right| > t_{n-k-1,\frac{\alpha}{2}}$ أي المعلم β_j له معنوية إحصائية فهو يختلف معنويا عن الصفر.

عندما يكون حجم العينة كبيرا ($n > 30$) فينبغي استعمال التوزيع الطبيعي و يمكن أخذ القيمة الحرجة $z_{\alpha/2}$ و ذلك بحساب المساحة المظلة للتوزيع.

2.5. اختبار المعنوية الكلية للنموذج و اختبارات القيود على المعالم

يمكن اختبار المعنوية الإجمالية للنموذج باستخدام نسبة التباين المفسر، إلى التباين غير المفسر، ويتبع هذا توزيع فيشر F، بدرجات حرية k و $n-k-1$، حيث n عدد المشاهدات و $k+1$ عدد المعالم المقدرة:

$$H_0 : \beta_0 = \beta_1 = ... = \beta_j = ... = \beta_k = 0$$

ضد الفرضية البديلة: $H_1 : \exists$ معامل $\neq 0$

$$F_c = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \overline{Y})^2 \Big/ k}{\sum_{i=1}^{n}\hat{\varepsilon}_i^2 \Big/ (n-k-1)} = \frac{\sum_{i=1}^{n}\hat{y}_i^{\,2} \Big/ k}{\sum_{i=1}^{n}\hat{\varepsilon}_i^2 \Big/ (n-k-1)} = \frac{R^2/k}{(1-R^2)/(n-k-1)} \sim F_\alpha(k, n-k-1)$$

فإذا تجاوزت الإحصائية F قيمة F المجدولة عند مستوى معنوية α وبدرجتي حرية k و $n-k-1$ نقبل الفرضية القائلة بأن معالم النموذج ليست جميعها مساوية للصفر وأن R^2 يختلف جوهريا عن الصفر. في هذه الحالة، يمكن القول أن للنموذج معنوية إحصائية.

هناك اختبارات أخرى تعتمد على جدول تحليل التباين (إدخال متغير أو عدة متغيرات مفسرة إضافية، استقرار معاملات النموذج، اختبار القيود على المعاملات..، الخ):

$$H_0 : R\beta = r$$

$$H_1 : R\beta \neq r$$

$$F_c = \frac{\left\{(R\hat{\beta} - r)'\left[R(X'X)^{-1}R'\right]^{-1}(R\hat{\beta} - r)\right\}/q}{RSS/(n-k-1)}$$

حيث $\hat{\beta}$ شعاع المعالم المقدرة للنموذج غير المقيد. نرفض H_0 إذا كانت F_c أكبر من القيمة المجدولة لتوزيع فيشر بدرجتي حرية q و $n-k-1$ و بطريقة أخرى، يمكن استعمال الإحصائية التالية:

$$F_c = \frac{(RSS_c - RSS_{nc})/q}{RSS_{nc}/(n-k-1)}$$

حيث RSS_{nc} مجموع مربعات بواقي تقدير النموذج غير المقيد و RSS_c الخاص بالنموذج المقيد.

هناك اختبار آخر مكافئ لاختبار فيشر يرتكز على مقارنة نسبة المعقولية للنموذج المقيد و غير المقيد. إذا كانت القيود موجودة هذا يعني أن $L_c < L_{nc}$ حيث L_{nc} هي دالة المعقولية للنموذج غير المقيد و L_c للنموذج المقيد، أي أن $L_c / L_{nc} < 1$ أو بشكله اللوغاريتمي $\ln L_c - \ln L_{nc} < 0$. الفرق بين لوغاريتمات الدالة ينبغي أن يكون معنويا سالبا. يمكن أن نبرهن أن هذا الاختبار يقودنا إلى اختبار χ^2 و ذلك بحساب الإحصائية $LR = -2(\ln L_c - \ln L_{nc})$ الذي تتبع بطبيعة الحال توزيع χ^2 بدرجة حرية r و التي تعبر عن عدد القيود. إضافة إلى ذلك، إذا كان LR أكبر من القيمة المجدولة لتوزيع χ^2 بنسبة معنوية α و درجة حرية r، نرفض الفرضية H_0 أي أن القيود ليست محققة. كما أنه يمكننا استعمال مضاعف لاغرانج[1]

مثال 3:

بالاستعانة بمعطيات المثال السابق، نقوم باختبار المعنوية الإحصائية لمعالم النموذج. نتساءل ما إذا كان يمكن قبول هذه المقدرات كأساس للوصول إلى معالم المجتمع الإحصائي.

نبني أولا مجالات ثقة للمعالم:

$$\beta_0 \in \left[\hat{\beta}_0 - t_{0.025}\hat{\sigma}_{\hat{\beta}_0}, \hat{\beta}_0 + t_{0.025}\hat{\sigma}_{\hat{\beta}_0}\right]$$
$$\beta_1 \in \left[\hat{\beta}_1 - t_{0.025}\hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + t_{0.025}\hat{\sigma}_{\hat{\beta}_1}\right]$$

1- سنتطرق إلى هذا الاختبار في الفصل الثالث بغية اختبار وجود الارتباط الذاتي و تجانس التباين.

$$\beta_2 \in \left[\hat{\beta}_2 - t_{0.025}\hat{\sigma}_{\hat{\beta}_2} , \hat{\beta}_2 + t_{0.025}\hat{\sigma}_{\hat{\beta}_2}\right]$$

حيث $t_{0.025}$ هي القيمة الحرجة لتوزيع ستيودنت بنسبة معنوية 5% و درجة حرية $n-2-1=7$ والتي تعادل القيمة 2.365 في جدول توزيع ستيودنت. بالتطبيق العددي لدينا:

$$\beta_0 \in [1.0996 - 2.365 \times 0.5513 , 1.0996 + 2.365 \times 0.5513]$$

أي: $\beta_0 \in [-0.2042 \ , \ 2.4034]$

و:

$$\beta_1 \in [0.9776 - 2.365 \times 0.0858 \ , \ 0.9776 + 2.365 \times 0.0858]$$

هذا يعني: $\beta_1 \in [0.7746 \ , \ 1.1805]$

أما :

$$\beta_2 \in [0.1237 - 2.365 \times 0.1190 \ , \ 0.1237 + 2.365 \times 0.1190]$$

و منه: $\beta_2 \in [-0.1577 \ , \ 0.4051]$

نقوم الآن بالاختبار. لدينا الفرضيتان:

$$H_0 : \beta_1 = 0$$

ضد: $H_1 : \beta_1 \neq 0$

نعلم أن $\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}}$ تتبع توزيع ستيودنت بدرجة حرية $n-3$ و في ظل قبول الفرضية H_0 :

$t_c = \frac{\hat{\beta}_1 - 0}{\hat{\sigma}_{\hat{\beta}_1}} = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}$ تتبع توزيع ستيودنت بدرجة حرية $n-3$.

لدينا إذن: $t_c = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}} = \frac{0.9776}{0.0858} = 11.3939$

نتخذ القرار وذلك بمقارنة القيمة المطلقة للقيمة المحسوبة بالقيمة الحرجة لتوزيع ستيودنت بدرجة حرية 7 و نسبة معنوية 5%. نلاحظ أن $|t_c| = 11.3939 > t_{0.025} = 2.365$ و عليه

نقبل الفرضية البديلة H_1 أي أن β_1 يختلف معنويا عن الصفر بنسبة معنوية $\alpha = 5\%$ وهذا يعني أنه يمكن قبول المقدر كأساس للوصول إلى معلمه النظري في المجتمع الإحصائي.

أما المعلم β_2، فلدينا:

$$H_0 : \beta_2 = 0$$

ضد: $H_1 : \beta_2 \neq 0$

نعلم أن $\frac{\hat{\beta}_2 - \beta_2}{\hat{\sigma}_{\hat{\beta}_2}}$ تتبع توزيع ستيودنت بدرجة حرية $n-3$ و في ظل قبول الفرضية H_0 :

$t_c = \frac{\hat{\beta}_2 - 0}{\hat{\sigma}_{\hat{\beta}_2}} = \frac{\hat{\beta}_2}{\hat{\sigma}_{\hat{\beta}_2}}$ تتبع توزيع ستيودنت بدرجة حرية $n-3$.

إذن:

$$t_c = \frac{\hat{\beta}_2}{\hat{\sigma}_{\hat{\beta}_2}} = \frac{0.1237}{0.1190} = 1.0394$$

نلاحظ في هذه الحالة أن $|t_c| = 1.0394 < t_{0.025} = 2.365$ و عليه نرفض الفرضية البديلة H_1 أي أن β_2 يساوي معنويا الصفر بنسبة معنوية $\alpha = 5\%$ وهذا يعني أنه لا يمكن قبول هذا المقدر كأساس للوصول إلى معلمه النظري في المجتمع الإحصائي.

نختبر الآن المعنوية الكلية للنموذج معتمدا على إحصائية فيشر. لدينا الفرضيتان:

$$H_0 : \beta_0 = \beta_1 = \beta_2 = 0$$

ضد: $H_1 : \exists \beta_j \neq 0, j = 1,2$

نعلم أيضا أن $\frac{R^2/k}{(1-R^2)/(n-k-1)}$ تتبع توزيع فيشر بدرجتي حرية k و $n-k-1$.

لدينا في هذه الحالة:

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{0.9929/2}{(1-0.9929)/(10-2-1)} = 490.7224$$

يتم اتخاذ القرار بمقارنة القيمة المحسوبة بالقيمة الحرجة لتوزيع فيشر بدرجتي حرية 2 و 7 و نسبة معنوية 5%. نلاحظ أن $F = 490.7224 > F_{0.05}(2,7) = 4.74$ و عليه نقبل الفرضية البديلة H_1 أي أن للنموذج معنوية إحصائية بنسبة معنوية $\alpha = 5\%$.

إذا اعتبرنا أن النموذج المراد دراسته هو $Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i$ ، هل إضافة المتغير X_{i2} سيحسن معنويا جودة التقدير؟ بعبارة أخرى، أي النموذج نختار؟

نقوم أولا بحساب مجموع المربعات الكلية TSS ، مجموع المربعات المقدرة ESS و مجموع مربعات البواقي RSS للنموذج الأصلي (ذي متغيرين) غير المقيد. من خلال المثال 2، تحصلنا على هذه النتائج:

$$TSS = \sum_{i=1}^{10}\left(Y_i - \overline{Y}\right)^2 = 256.4$$

$$ESS = \sum_{i=1}^{10}\left(\hat{Y}_i - \overline{Y}\right)^2 = 254.58$$

$$RSS = \sum_{i=1}^{10}\hat{\varepsilon}_i^2 = 1.815$$

ثانيا، نقوم بحساب مجموع المربعات الكلية TSS_c ، مجموع المربعات المقدرة ESS_c و مجموع مربعات البواقي RSS_c للنموذج ذي متغير واحد (المقيد). النموذج المقدر هو كالتالي:

$$\hat{Y}_i = 1.53 + 1.05X_i$$
$$(4.24) \quad (31.15)$$
$$R^2 = 0.9918$$
$$RSS_c = 2.095$$

نحسب إذن:

$$TSS_c = \frac{RSS_c}{1-R^2} = \frac{2.095}{1-0.9918} = 255.48$$
$$ESS_c = TSS_c - RSS_c = 255.48 - 2.095 = 253.39$$

يكون وضع الاختبار على الشكل التالي:

$$H_0 : \beta_2 = 0$$
$$H_1 : \beta_2 \neq 0$$

عندما نضيف متغيرات مستقلة جديدة في النموذج فسوف يؤدي إلى زيادة في ESS أي انخفاض في RSS. نود إذن اختبار أن الفرق بين RSS_c و RSS معنويا موجب. أي:

$$F_c = \frac{(RSS_c - RSS)/(k - k')}{RSS/(n-k-1)} = \frac{(2.095 - 1.815)/(2-1)}{1.815/(10-2-1)} = 1.0802$$

مع k: عدد المتغيرات المستقلة في النموذج غير المقيد و k' عدد المتغيرات المستقلة في النموذج المقيد (أي بدون إضافة المتغيرات المستقلة الأخرى). نقبل الفرضية H_0 ($F_c < F_{0.05}(1,7)$) أي ليس هناك فرق معنوي بين تبايني البواقي، إضافة المتغير المستقل X_{i2} لا يحسن بصفة معنوية القدرة التفسيرية للنموذج.

3.5. اختبار استقرار معاملات النموذج – اختبار Chow –

يدرس هذا الاختبار مدى استقرار النموذج في كامل الفترة الزمنية (دراسة التغيير الهيكلي للنموذج)، أي صياغة النموذج هي نفسها ولكن تختلف القيم المقدرة للمعاملات في العينتين الجزئيتين. ليكن النموذج المقدر ذو k متغير مستقل على فترة واحدة:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \ldots + \hat{\beta}_k X_{ik}$$ على عينة حجمها n

نقدر النموذج انطلاقا من عينتين جزئيتين n_1 و n_2 مع $n = n_1 + n_2$، حيث:

$$\hat{Y}_i = \hat{\beta}_0^{(1)} + \hat{\beta}_1^{(1)} X_{i1} + \hat{\beta}_2^{(1)} X_{i2} + \ldots + \hat{\beta}_k^{(1)} X_{ik}$$
$$\hat{Y}_i = \hat{\beta}_0^{(2)} + \hat{\beta}_1^{(2)} X_{i1} + \hat{\beta}_2^{(2)} X_{i2} + \ldots + \hat{\beta}_k^{(2)} X_{ik}$$

نختبر الفرضيات التالية:

$$H_0:\begin{pmatrix} \beta_0=\beta_0^{(1)}=\beta_0^{(2)} \\ \beta_1=\beta_1^{(1)}=\beta_1^{(2)} \\ \beta_2=\beta_2^{(1)}=\beta_2^{(2)} \\ \dots\dots\dots\dots\dots\dots \\ \beta_k=\beta_k^{(1)}=\beta_k^{(2)} \end{pmatrix}$$

إن اختبار استقرار المعاملات يقودنا الى طرح السؤال التالي: هل يوجد فرق معنوي بين مجموع مربعات البواقي في كامل الفترة n وجمع مجموع مربعات البواقي المحسوبة انطلاقا من العينتين الجزئيتين RSS^1+RSS^2؟ إذا كانت الإجابة "لا"، فهذا يعني أن النموذج مستقر في كامل العينة.

تعرف احصائية فيشر كما يلي:

$$F_c=\frac{\left[RSS-\left(RSS^1+RSS^2\right)\right]/df_1}{\left(RSS^1+RSS^2\right)/df_2}$$

مع:

$$df_1=(n-k-1)-[(n_1-k-1)+(n_2-k-1)]=k+1$$
$$df_2=(n_1-k-1)+(n_2-k-1)=n-2(k+1)$$

إذا كانت $F_c\le F_\alpha(k+1,n-2(k+1))$، ففي هذه الحالة نقبل الفرضية H_0، أي أن المعاملات مستقرة معنويا في كامل الفترة الزمنية.

مثال 4:

لنعتبر النموذج المقدر التالي على عينة حجمها $n=14$ [1]:

$$\hat{Y}_i=32.89+0.80X_{i1}-0.38X_{i2}-0.03X_{i3}$$
$$(11.66)\quad(0.29)\quad(0.15)\quad(0.05)$$
$$R^2=0.702$$
$$n=14$$

القيم التي بين قوسين هي قيم الأخطاء المعيارية.

1- Bourbonnais (2003), p. 70

الخطوات هي كالتالي:

الخطوة الأولى: تقدير النموذج على عينتين جزئيتين وحساب مجموع مربعات البواقي

الفترة الجزئية الأولى:

$$\hat{Y}_i = 25.27 + 0.774X_{i1} - 0.293X_{i2} - 0.012X_{i3}$$
$$(0.53) \qquad (0.31) \qquad (0.10)$$

$R^2 = 0.692$

$n_1 = 7$

$TSS^1 = 88.85;\ ESS^1 = 61.54\ ;\ RSS^1 = 27.31$

الفترة الجزئية الثانية:

$$\hat{Y}_i = 62.63 + 1.228X_{i1} - 0.62X_{i2} - 0.184X_{i3}$$
$$(0.69) \qquad (0.52) \qquad (0.15)$$

$R^2 = 0.543$

$n_2 = 7$

$TSS^2 = 45.43;\ ESS^2 = 24.70\ ;\ RSS^2 = 20.73$

الخطوة الثانية: حساب إحصائية فيشر

$$F_c = \frac{\left[RSS - \left(RSS^1 + RSS^2\right)\right]/(k+1)}{\left(RSS^1 + RSS^2\right)/\left(n - 2(k+1)\right)} = \frac{\left[67.45 - (27.31 + 20.73)\right]/4}{(27.31 + 20.73)/6} = \frac{4.852}{8.00}$$
$$= 0.60 < F_{0.05}(4,6) = 4.53$$

نقبل الفرضية H_0، يمكن القول أن معاملات النموذج مستقرة معنويا على كامل الفترة الزمنية بنسبة مخاطرة 5%.

6. استخدام المتغيرات الصورية

في كثير من الأحيان، نعتقد بوجود متغيرات معينة ذات أهمية عظمى ولكن من طبيعة نوعية لا يأخذ إلا القيمتين 0 أو 1. نستعمل هذا النوع من المتغيرات عندما نريد دمج عامل مستقل ثنائي: "الظاهرة حدثت أو لم تحدث" أو أيضا عندما يكون العامل المستقل ذا طابع نوعي: "ذكر أو أنثى"...

الظاهرة تحدث:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_i$$

الظاهرة لا تحدث:

$$Y_i = \alpha_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_i$$

يمكن كتابة هاتين المعادلتين في معادلة واحدة:

$$Y_i = \beta_0 + (\alpha_0 - \beta_0) D_i + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_i$$

مع $D_i = 1$ عندما لا تحدث الظاهرة

و $D_i = 0$ عندما تحدث الظاهرة.

في هذه الحالة نقوم بإدماج متغير مفسر إضافي للنموذج المحدد الأصلي و تطبيق الطرق الكلاسيكية للتقدير ويطلق على المتغير D_i الذي يظهر في المعادلة المتغير الصوري Dummy variable، وهو كما ذكرنا يأخذ قيمة الواحد الصحيح إذا تحققت بعض الشروط و قيمة الصفر إذا لم تتحقق وعليه فإن استخدام المتغيرات الصورية تعميم قوي لتحليل الانحدار، فإنه يسمح لنا تعميم مجال تحليلنا للانحدار ليشمل متغيرات مهمة لا يمكننا وصفها في وحدات كمية. فباستخدام هذا النوع من المتغيرات يمكن الأخذ في الحسبان تأثير العوامل النوعية المهمة التي تؤثر على المتغير التابع.

مثال 5:

سنأخذ مثالا يتضمن نوعا من المشاكل يتمثل في التقلبات الناتجة عن تأثير حرب الخليج. بعد ما تأكدنا من أهمية القيمة المضافة للقطاع السياحي X_{t1} و عدد السكان X_{t2}، نتسأل ما إذا كان لحرب الخليج الثانية تأثير على الإنتاج السياحي Y_t. الطريقة إذن هي إدخال متغير صوري بقيمة الواحد إذا كنا في حالة الحرب أو قيمة الصفر في حالة السلم، حيث:

$D_i = 0$ إذا كانت عدد المشاهدات يتراوح بين 1975 إلى غاية 1990 و أيضا 1992.

$D_i = 1$ في سنة 1991.

النموذج المقدر يكون على الشكل التالي:

$$\hat{Y}_t = 2.340.4 + 23.5X_{t1} + 0.3X_{t2} - 120.56D_i$$
$$(4.5) \quad (2.2) \quad (2.9) \quad (5.8)$$
$$n = 18$$
$$R^2 = 0.65$$

حيث (.): قيم ستيودنت.

نلاحظ أن $|t_c| = 5.8 > t_{0.025} = 2.14$ و عليه نقبل الفرضية البديلة H_1 أي أن معامل المتغير الصوري يختلف معنويا عن الصفر بنسبة معنوية $\alpha = 5\%$ وهذا يعني أن الإنتاج السياحي في سنة 1991 منخفض جدا (120.56) و هذا الانخفاض ناتج عن أثر حرب الخليج.

7. التنبؤ العلمي باستعمال الانحدار الخطي المتعدد

نتطرق في هذه الفقرة إلى قضية التنبؤ بالملاحظات المستقبلية، (أو خارج العينة) لشعاع الملاحظات الخاصة بالمتغير التابع $Y_t : t = 1,2,...,T$، وذلك بمعرفتنا لمصفوفة الملاحظات المستقبلية للمتغيرات المستقلة، فليكن النموذج الخطي العام خلال العينة T والمقدر على الشكل: $\hat{Y} = X\hat{\beta}$

ومقدر المربعات الصغرى العادية $\hat{\beta} = (X'X)^{-1}X'Y = AY$ ، ويكون المقدر بملاحظة واحدة في المستقبل هو:

$$\hat{Y}_T(1) = \hat{\beta}_0 + \hat{\beta}_1 X_{T+1,1} + \hat{\beta}_2 X_{T+1,2} + \hat{\beta}_3 X_{T+1,3} + ... + \hat{\beta}_k X_{T+1,k}$$

التنبؤ بعد فترتين في المستقبل:

$$\hat{Y}_T(2) = \hat{\beta}_0 + \hat{\beta}_1 X_{T+2,1} + \hat{\beta}_2 X_{T+2,2} + \hat{\beta}_3 X_{T+2,3} + ... + \hat{\beta}_k X_{T+2,k}$$

التنبؤ بعد الفترة h في المستقبل:

$$\hat{Y}_T(h) = \hat{\beta}_0 + \hat{\beta}_1 X_{T+h,1} + \hat{\beta}_2 X_{T+h,2} + \hat{\beta}_3 X_{T+h,3} + ... + \hat{\beta}_k X_{T+h,k}$$

حيث $h=1,2,....,H$ يسمى بأفق التنبؤ.

وعليه نصل إلى التنبؤ بالفترة H في المستقبل:

$$\hat{Y}_T(H)=\hat{\beta}_0+\hat{\beta}_1 X_{T+H,1}+\hat{\beta}_2 X_{T+H,2}+\hat{\beta}_3 X_{T+H,3}+..........+\hat{\beta}_k X_{T+H,k}$$

إذن إذا أردنا التنبؤ بمجموعة من الملاحظات المستقبلية بفترة تساوي H ملاحظة مرة واحدة يكون شعاع القيم التقديرية:

$$\hat{Y}_T(H)=\underset{(H\times 1)}{\begin{pmatrix}\hat{Y}_T(1)\\ \hat{Y}_T(2)\\ \vdots\\ \vdots\\ \hat{Y}_T(H)\end{pmatrix}}$$

أما مصفوفة ملاحظة المتغيرات المستقلة المستقبلية فهي:

$$X_{T+H}=\underset{(H\times(k+1))}{\begin{pmatrix}1\ X_{T+1,1} & X_{T+1,2} & \cdots & X_{T+1,k}\\ 1\ X_{T+2,1} & X_{T+2,2} & \cdots & X_{T+2,k}\\ \vdots\ \ \vdots & \vdots & \vdots & \vdots\\ \vdots\ \ \vdots & \vdots & \ddots & \vdots\\ 1\ X_{T+H,1} & X_{T+H,2} & \cdots & X_{T+H,k}\end{pmatrix}}$$

ومنه يمكن كتابة النموذج الخطي العام المتنبأ به على الشكل: $Y_{T+H}=X_{T+H}\beta+\varepsilon_{T+H}$، كما يكون النموذج المقدر من الشكل: $\hat{Y}_T(H)=X_{T+H}\hat{\beta}$، ويكون متوسط مقدر التنبؤ هو:

$$E\left(\hat{Y}_T(H)\right)=X_{T+H}E\left(\hat{\beta}\right)=X_{T+H}\beta=E\left(Y_{T+H}\right)$$

ومنه نقول أن $\hat{Y}_T(H)$ هو تنبؤ خطي غير متحيز للعبارة:

$$X_{T+H}\beta=E\left(Y_{T+H}\right)$$

ليكون التباين:

$$\operatorname{var}\left(\hat{Y}_T(H)\right)=\left[\left(\hat{Y}_T(H)-X_{T+H}\beta\right)\left(\hat{Y}_T(H)-X_{T+H}\beta\right)'\right]$$
$$=\sigma_\varepsilon^2 X_{T+H}\left(X'X\right)^{-1}X'_{T+H}$$

لنعرف شعاع أخطاء التنبؤ:

$$\hat{\varepsilon}_{T+H}=Y_{T+H}-\hat{Y}_T(H)$$
$$E\left(\hat{\varepsilon}_{T+H}\right)=E\left(Y_{T+H}-\hat{Y}_T(H)\right)=0$$

أما تباين شعاع أخطاء التنبؤ فهو:

$$\operatorname{var}\left(\hat{\varepsilon}_{T+H}\right)=\operatorname{var}\left(Y_{T+H}-\hat{Y}_T(H)\right)=E\left[\left(-X_{T+H}\left(\hat{\beta}-\beta\right)+\varepsilon_{T+H}\right)\left(-X_{T+H}\left(\hat{\beta}-\beta\right)+\varepsilon_{T+H}\right)'\right]$$

لنجد في الأخير:

$$\operatorname{var}\left(\hat{\varepsilon}_{T+H}\right)=\sigma_\varepsilon^2 X'_{T+H}\left(X'X\right)^{-1}X_{T+H}+\sigma_\varepsilon^2 I_H$$

ويكون هذا التنبؤ هو أحسن تنبؤ خطي غير متحيز (BLUP) يمكن الحصول عليه، حيث إذا عرفنا $\widetilde{Y}_T(H)$ تنبؤ آخر خطي لعينة ملاحظات المتغير التابع مع متوسط خطا التنبؤ مساو للصفر $E\left(\widetilde{\varepsilon}_{T+H}\right)=E\left(Y_{T+H}-\widetilde{Y}_T(H)\right)=0$ تكون لدينا المتراجحة:

$$\operatorname{var}\left(Y_{T+H}-\widetilde{Y}_T(H)\right)-\operatorname{var}\left(Y_{T+H}-\hat{Y}_T(H)\right)\geq 0$$

ومنه يمكن القول أن $\hat{Y}_T(H)=X_{T+H}\hat{\beta}$ هو أحسن تنبؤ خطي غير متحيز.

وتكون اختبارات التنبؤ عن طريق إيجاد التوزيع الذي يعتبر فرضية العدم، والقائلة بأن النموذج الخطي العام يبقى محافظا على شكله من الملاحظة الأولى إلى الملاحظة $T+H$ في المستقبل، أي نفترض عدم تغير البناء الهيكلي للنموذج،

$$H_0:\quad \hat{Y}=X\hat{\beta}\quad t=1,2,3,\ldots\ldots T,T+1,\ldots,T+h,\ldots.,T+H$$

وذلك ضد الفرضية البديلة، والتي تقول أن نموذج العينة الأولى T يختلف عن نموذج التنبؤ للفترة H.

$$F=\frac{\left(Y_{T+H}-\hat{Y}_T(H)\right)'\left[X_{T+H}\left(X'X\right)^{-1}X'_{T+H}+I_H\right]^{-1}\left(Y_{T+H}-\hat{Y}_T(H)\right)/H}{\hat{\sigma}_\varepsilon^2} \quad \rightsquigarrow \quad F_{H,T-k-1}$$

وإذا كان $H=1$ يصبح التوزيع أعلاه على الشكل:

$$F=\frac{\left(Y_{T+1}-\hat{Y}_T(1)\right)'\left[X_{T+1}\left(X'X\right)^{-1}X'_{T+1}+1\right]^{-1}\left(Y_{T+1}-\hat{Y}_T(1)\right)}{\hat{\sigma}_\varepsilon^2} \quad \rightsquigarrow \quad F_{1,T-k-1}$$

مثال 6:

نستعمل دائما معطيات المثال الأول للتنبؤ بالمتغير Y للفترة 11 ثم بناء فترات ثقة بنسبة معنوية 5%.

إذا علمنا أن القيمة المستقبلية لكل من X_1 و X_2 هما على الترتيب 18 و 16 فإنه يمكن التنبؤ بالظاهرة Y انطلاقا من النموذج المقدر:

$$\hat{Y}_{11}=\hat{Y}_{10}(1)=1.0996+0.9776X_{11,1}+0.1237X_{11,2}$$
$$=1.0996+0.9776\times 18+0.1237\times 16=20.6756$$

لإيجاد مجال الثقة للتنبؤ، فلا بد من حساب الانحراف المعياري لخطأ التنبؤ. لدينا:

$$\hat{\sigma}^2_{\hat{\varepsilon}_{T+H}}=\hat{\sigma}^2_\varepsilon\left(X'_{T+H}\left(X'X\right)^{-1}X_{T+H}+I_H\right)$$
$$=(0.2593)\times\left(\begin{pmatrix}1 & 18 & 16\end{pmatrix}\begin{pmatrix}1.1725 & 0.0852 & -0.1921\\ 0.0852 & 0.0284 & -0.0362\\ -0.1921 & -0.0362 & 0.0547\end{pmatrix}\begin{pmatrix}1\\ 18\\ 16\end{pmatrix}+1\right)$$

و منه تباين خطأ التنبؤ: $\hat{\sigma}^2_{\hat{\varepsilon}_{T+1}}=0.2593(0.4461+1)=0.3749$

نستنتج الانحراف المعياري لخطأ التنبؤ: $\hat{\sigma}_{\hat{\varepsilon}_{T+1}}=0.6123$

إذن مجال الثقة يكون كما يلي:

$$Y_{T+1}\in\left[\hat{Y}_T(1)-t_{0.025}\hat{\sigma}_{\hat{\varepsilon}_{T+1}}\ ,\ \hat{Y}_T(1)-t_{0.025}\hat{\sigma}_{\hat{\varepsilon}_{T+1}}\right]$$

بالتطبيق العددي:

$$Y_{11}\in\left[20.6756-2.365\times 0.6123,\ 20.6756+2.365\times 0.6123\right]$$
$$Y_{11}\in\left[19.2275\ ,\ 22.1236\right]$$

الفصل الثالث

مشاكل القياس الاقتصادي

اختراق فرضيات النموذج

الفصل الثالث

مشاكل القياس الاقتصادي

اختراق فرضيات النموذج

يُغطي هذا الجزء ثلاثة مشاكل قياسية تواجه الباحث، تتعلق كل منها بإسقاط إحدى الفرضيات الكلاسيكية لطريقة المربعات الصغرى المذكورة آنفا وتتمثل هذه المشاكل في:

1. التعدد الخطي.
2. الارتباط الذاتي للأخطاء.
3. عدم ثبات تباين الخطأ.

1. التعدد (الازدواج) الخطي Multicollinearity:

إحدى فرضيات النموذج الكلاسيكي للانحدار المتعدد هي أن لمصفوفة المشاهدات عن المتغيرات المستقلة رتبة تامة k، هذه الفرضية سمحت لنا باستنتاج مقدر $\hat{\beta}$ لـ β ، خطي وغير متحيز وذي تشتت أصغر، وذلك انطلاقا من المعادلة $(X'X)\hat{\beta} = X'Y$. فإذا رفعت هذه الفرضية، فإن $(X'X)$ لن تكون ذات رتبة تامة، أي تكون أقل من رتبة(X) (أو (X')) أي أقل من k. ومع أن $(X'X)$هي مصفوفة ذات بعد ($k \times k$) بالتالي تكون مصفوفة شاذة (محددها معدوم)، ومنه فإن $(X'X)^{-1}$ تكون غير موجودة وعليه لا تقبل المعادلة $(X'X)\hat{\beta} = X'Y$ حلا وحيدا (عدد لا نهائي من الحلول). يضع النموذج الكلاسيكي للانحدار المتعدد $Y = X\beta + \varepsilon$ المتغير التابع $Y_i : i = 1....n$ في علاقة خطية مع المتغيرات المستقلة $X_{i1}, X_{i2},......X_{ik} : i = 1....n$، وكذلك مع الأخطاء العشوائية $\varepsilon_i : i = 1....n$ ، فإذا كانت بالإضافة إلى ذلك رتبة X أقل من أو تساوي k فإن هذا يترجم بارتباط خطي بين أعمدة المصفوفة X.

وبعبارة أخرى يشير مشكل التعدد الخطي إلى وجود ارتباط خطي بين عدد من المتغيرات المفسِّرة، ومن ثم فإن هذا المشكل لا يوجد في حالة الانحدار البسيط[1]. نسمي X_j العمود رقم j لـ X. حيث $X = \lfloor X_1, X_2, ...X_j .....X_k \rfloor$.

- قولنا أن رتبة X أقل من k يعني أنه يوجد شعاع C حيث:

$$C' = \lfloor C_1, C_2, ...C_j .....C_k \rfloor \neq 0$$

مع: $$C_1X_1 + C_2X_2 + ... + C_kX_k = 0$$

- العلاقة الأخيرة تعبر عن وجود علاقة خطية بين المتغيرات المستقلة.

1.1. أسباب التعدد الخطي وآثاره:

ينشأ التعدد الخطي من عدة أسباب منها ما يلي:

- ❖ اتجاه المتغيرات الاقتصادية معا للتغير مع مرور الزمن: فبمرور الزمن سوف تتزايد المتغيرات الاقتصادية التالية معا: الدخل، الاستهلاك، الادخار، الاستثمار، المستوى العام للأسعار والعمالة، وحيث أن هناك ارتباط بين هذه المتغيرات فإن التعدد الخطي سوف يتحقق.
- ❖ استخدام متغيرات مستقلة ذات فترة إبطاء في المعادلة المراد تقديرها: فالدخل في الفترة الزمنية الحالية يتحدد جزئيا بواسطة قيمته في الفترة الزمنية السابقة، وحيث أن هناك ارتباط بين القيم المتتالية لمتغير ما فإن التعدد الخطي سوف يتحقق.

وفي وجود التعدد الخطي فإنه سوف يترتب عنه:

- ❖ زيادة التباين والتباين المشترك للمقدرات بدرجة كبيرة دون التأثير على التنبؤات المستمدة من الانحدار[2].
- ❖ القيم المقدرة لمعاملات الانحدار سوف تكون غير محددة وغير دقيقة.
- ❖ الأخطاء المعيارية للقيم المقدرة لمعاملات الانحدار سوف تكون كبيرة جدا.

1- عبد القادر محمد عبد القادر عطية، 1990، ص 410.
2- امتثال محمد حسن، محمد علي محمد أحمد،، 2000، ص 354.

2.1. اختبارات اكتشاف التعدد الخطي:

تعتمد درجة الخطورة لأثر التعدد الخطي على درجة الارتباط الجزئي، ومعامل الارتباط الكلي (أو معامل التحديد المضاعف)، ومنه يمكن القول بأن كلا من الأخطاء المعيارية ومعاملات الارتباط الجزئية $r_{xi,xj}$، معامل التحديد المضاعف R^2، يمكنها أن تستعمل لاختبار التعدد الخطي، لكن كل معيار من هذه المعايير الثلاثة المذكورة ليس بمؤشر على وجود التعدد الخطي بمفرده، وذلك لأن القيم العالية للأخطاء المعيارية لا تظهر دائما، بسبب التعدد الخطي، وإنما يمكن أن تظهر لأسباب أخرى، كما أن الارتباطات العالية فيما بين المتغيرات المستقلة لا تؤثر بالضرورة على قيم المقدرات $\hat{\beta}_j$، ومنه ليست هذه الأخيرة بمعيار مناسب لقياس واكتشاف التعدد الخطي بمفردها، وبالمقابل يمكن لقيمة معامل التحديد المضاعف R^2 أن تكون عالية بالمقارنة مع $r_{Xi,Xj}$.

ورغم ذلك، من المحتمل أن تحتوي نتائجنا على إشارات خاطئة أو على أخطاء معيارية كبيرة، ومع كل هذا يمكن القول بأن توفيق المعايير الثلاثة، أعلاه يساعدنا على اكتشاف التعدد الخطي.

1.2.1. طريقة التحليل الترافدي لـ Frisch :

تكمن هذه الطريقة في انحدار المتغير التابع على كل متغير مستقل على حدا، ومنه نحصل على كل الانحدارات الأولية، ثم نختار الانحدار الأولي الذي يعطي النتائج الأكثر مصداقية، ثم نضيف تدريجيا متغيرات أخرى ونختبر آثارها على كل من المعالم الفردية (أخطائها المعيارية، قيمة R^2) ويكون المتغير المضاف للانحدار ذا معنوية إذا تحققت فيه الشروط التالية:

- ❖ إذا حَسَّن المتغير المستقل الجديد من R^2 بدون أن يجعل المعالم الفردية مرفوضة بطريقة خاطئة، نحتفظ بهذا المتغير ونعتبره كمتغير مستقل.
- ❖ إذا لم يُحَسِّن المتغير الجديد من العلاقة ويؤثر على قيم المعالم الفردية، نعتبره مرفوضا ونحذفه من الانحدار.

❖ إذا أثر المتغير الجديد بشكل واضح على إشارات وقيم المعالم المقدرة، نعتبره متغيرا مُفَسِّرا، فإذا تأثرت المعالم الفردية بالطريقة التي تصبح فيها غير مقبولة على أساس الاعتبارات النظرية المعروفة مسبقا، فإنه يمكننا القول بأن هذا مؤشر على وجود التعدد الخطي بشكل معقد، يكون هذا المتغير مُهما، لكن بسبب الارتباطات الخطية مع المتغيرات المستقلة الأخرى، يكون أثره غير مقدر وغير معروف إحصائيا بواسطة المربعات الصغرى العادية.

إن التحليل الترافدي لـ Frisch . ينص على تقدير كل الانحدارات الممكنة ما بين المتغيرات الموجودة بالعلاقة المدروسة، آخذين كل متغير، بالترتيب، كمتغير تابع واعتبار كل الانحدارات الممكنة لكل متغير في بقية المتغيرات، والتي ندخلها تدريجيا في التحليل، ومن الواضح أن التحليل الترافدي يتطلب منا حسابات كثيرة، ومنه تكون المقارنات ما بين النتائج معقدة أكثر.

2.2.1. قياس التعدد الخطي أو شرط الأعداد Condition numbers:

من خلال النموذج التالي: $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i : \quad i = 1 ....n$

يكون لدينا :

$$\begin{cases} Var(\hat{\beta}_1) = \dfrac{\sigma_\varepsilon^2}{\sum X_{i1}^2 (1 - R_1^2)} \\ Var(\hat{\beta}_2) = \dfrac{\sigma_\varepsilon^2}{\sum X_{i2}^2 (1 - R_2^2)} \\ Cov(\hat{\beta}_1, \hat{\beta}_2) = \dfrac{\sigma_\varepsilon^2 R_1^2}{\sum X_{i1} X_{i2} (1 - R_1^2)} \end{cases}$$

حيث أن R_1^2 هو مربع معامل الارتباط المتعدد ما بين المتغيرين المستقلين X_{i1} و X_{i2}، بينما R_2^2 هو ما بين $X_{i1} X_{i2}$، وهما في الأخير متساويان، أما عند توسيع النموذج إلى k متغير مستقل ($k > 2$) يصبح R_j^2 على أنه مربع معامل الارتباط المتعدد ما بين المتغير المستقل X_{ij} وبقية المغيرات المستقلة الأخرى، ومنه يمكننا استنتاج قانون عام لتباين المقدرات الفردية لشعاع معالم النموذج كما يلي:

$$Var(\hat{\beta}_j) = \frac{\sigma_\varepsilon^2}{\sum X_{ij}^2 (1 - R_j^2)}, \quad j = 1 ....k$$

وتكون قيمة $Var(\hat{\beta}_j)$ كبيرة كلما كانت: σ_ε^2 كبيرة؛ $\sum X_{ij}^2$ صغيرة، R_j^2 كبيرة.

ومنه نُعرِّف مقياسا جديد يسمى " معامل تضخم التباين" Variance Inflation Factor (V.I.F)" ، ومقياسا آخر يسمى "شرط العدد Condition number"، وهما مقياسان يحددان درجة التعدد الخطي.

- ويعرف معامل تضخم التباين كما يلي:

$$V.I.F(\hat{\beta}_j) = \frac{1}{1 - R_j^2}$$

- وبناءا على هذا التعريف نستطيع كتابة:

$$Var(\hat{\beta}_j) = \frac{\sigma_\varepsilon^2}{\sum X_{ij}^2} \times V.I.F(\hat{\beta}_j), \quad j = 1......k$$

أي أن:

$$V.I.F(\hat{\beta}_j) = \frac{\sum X_{ij}^2}{\sigma_\varepsilon^2} \times Var(\hat{\beta}_j) \qquad j = 1......k$$

انطلاقا من الانتقادات الموجهة لمعامل الارتباط، يكون مقياس VIF غير كافٍ لتحديد التعدد الخطي، ومنه نذكر مقياس شرط الأعداد المذكور من طرف Welsch (1980) ، والذي يقيس حساسية مقدرات الانحدار للتغيرات الصغيرة في التباينات، ويعرف شرط الأعداد على أنه الجدر التربيعي لأكبر قيمة مقسمة على أصغر قيمة للقيم المميزة للمصفوفة $(X'X)$ وهو على الشكل: $K(X) = \frac{\sqrt{\lambda_{\max}}}{\sqrt{\lambda_{\min}}}$

فكلما كانت القيمة أعلاه أقرب إلى الواحد، كلما كان الشرط أفضل لعدم جدية التعدد الخطي، ومع هذا، فإن المقياسين المذكورين أعلاه ليسا كاملين، حيث القانون الخاص بـ . VIF ينظر إلى الارتباطات من خلال المتغيرات المستقلة فقط، وهذا ليس بالعامل الوحيد، كما أن شرط العدد يمكن أن يتغير بإعادة تحويل المتغيرات المستقلة، والتي ليست دائمة صحيحة، ويصلح المقياسان للاستعمال عند حذف بعض المتغيرات وفرض قيود على المعالم فقط في الحالات التي يكون فيها $R_j^2 \approx 1$، أو لما تكون القيمة المميزة

الصغيرة $\lambda_{\min}$ أقرب من الصفر. نقدر النموذج في هذه الحالة تبعا لبعض القيود المفروضة على معالمه، ويقترح Theil مقياسا أخر لقياس درجة الارتباط فيما بين المتغيرات ومنه درجة التعدد الخطي على الشكل: $m = R^2 - \sum_{j=1}^{k}\left(R^2 - R_{-j}^2\right)$

حيث أن R^2 هو معامل التحديد المضاعف المعروف من قبل، أما R_{-j}^2 فهو مربع معامل الارتباط المتعدد من انحدار y (المركزة) في $x_1, x_2, \ldots\ldots x_k$. مع حذف x_j، لكن إحدى عيوب هذه الطريقة هي أن m يمكن أن تكون سالبة مما يجعل التحليل أصعب، وهناك من يقترح طرقا معينة لحل مشكلة التعدد الخطي كإضافة حد ثابت لتباينات مقدرات المعالم قبل حل المعادلات الطبيعية للمربعات الصغرى.

3.2.1. طريقة Farrar-Glauber: لاكتشاف ظاهرة التعدد الخطي يتبع Farrar-Glauber الخطوات التالية:

- أولا: حساب محدد مصفوفة معاملات الارتباط بين المتغيرات المستقلة:

$$D = \begin{vmatrix} 1 & r_{X_1X_2} & r_{X_1X_3} & \cdots & r_{X_1X_k} \\ r_{X_2X_1} & 1 & r_{X_2X_3} & \cdots & r_{X_2X_k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{X_kX_k} & r_{X_kX_2} & r_{X_kX_3} & \cdots & 1 \end{vmatrix}$$

عندما تكون قيمة المحدد تقترب من الصفر، فإن هناك دليل على وجود تعدد خطي.

- ثانيا: نستعمل اختبار χ^2 وذلك بوضع الفرضيات التالية:

$H_0 : \mathrm{D} = 1$ (استقلال خطي)

$H_1 : \mathrm{D} < 1$ (ارتباط خطي)

إحصائية Farrar-Glauber (القيمة المحسوبة) تعرف كما يلي:

$$\chi^2{*} = -\left[n - 1 - \frac{1}{6}(2k+7)\right].\ln D$$

حيث n هو حجم العينة، k هو عدد المتغيرات المفسرة في النموذج و ln هو اللوغاريتم النبري.

فإذا كانت قيمة χ^{2*} أكبر تماما من القيمة المجدولة لتوزيع χ^2 بدرجة حرية $\frac{1}{2}k(k+1)$ و نسبة معنوية α، نقبل H_1 أي هناك تعدد خطي.

مثال 1:

الهدف هو دراسة قياسية للاستثمار Y_i بدلالة معدل الفائدة X_{i1}، الإنتاج الوطني الإجمالي X_{i2} و أسعار الاستهلاك X_{i3}. نريد اختبار وجود التعدد الخطي بين المتغيرات المستقلة. لتكن المعطيات التالية:

الجدول (1): المعطيات الإحصائية

X_{i3}	X_{i2}	X_{i1}	Y_i	السنة
82.54	873.4	5.16	133.3	1988
86.79	944.0	5.87	149.3	1989
91.45	992.7	5.95	144.2	1990
96.01	1077.6	4.88	166.4	1991
100.00	1185.9	4.50	195.0	1992
105.75	1326.4	6.44	229.8	1993
115.08	1434.2	7.83	228.7	1994
125.79	1549.2	6.25	206.1	1995
132.34	1718.0	5.50	257.9	1996
140.05	1918.3	5.46	324.1	1997
150.42	2163.9	7.46	386.6	1998
163.42	2417.8	10.28	423.0	1999
178.64	2633.1	11.77	402.3	2000
195.51	2937.7	13.42	471.5	2001
207.23	3057.5	11.02	421.9	2002

لاختبار التعدد الخطي، نستعمل اختبار Farrar-Glauber لهذا الغرض. نحسب أولا مصفوفة الارتباطات بين المتغيرات المستقلة C:

$$C = \begin{pmatrix} 1 & 0.99 & 0.87 \\ 0.99 & 1 & 0.87 \\ 0.87 & 0.87 & 1 \end{pmatrix}$$

نحسب محدد هذه المصفوفة D : $D = \begin{vmatrix} 1 & 0.99 & 0.87 \\ 0.99 & 1 & 0.87 \\ 0.87 & 0.87 & 1 \end{vmatrix} = 0.0199$

نحسب إحصائية Farrar-Glauber

$$\chi^2* = -\left[n-1-\frac{1}{6}(2k+7)\right].\ln D = -\left[15-1-\frac{1}{6}(2(3)+7)\right](-3.91) = 46.26$$

نلاحظ أن هذه الإحصائية أكبر تماما من القيمة المجدولة لتوزيع χ^2 بدرجة حرية $\frac{1}{2}k(k+1) = 6$ و عليه نرفض H_0. هناك تعدد خطي أي أن المتغيرات المستقلة مرتبطة فيما بينها.

3.1. الحلول المقترحة للتعدد الخطي:

عند وجود التعدد الخطي، فإن الحلول تكون مُعتمِدة على إمكانية إيجاد مصادر أخرى للبيانات وعلى أهمية العوامل التي تسببت في ظهورها، ثم على الهدف الذي من أجله نقوم بتقدير الدالة تحت الدراسة، فإذا لم يؤثر التعدد الخطي بشكل فعلي على مقدرات النموذج، يقترح بعض باحثي القياس الاقتصادي إهمال وجوده في النموذج، حيث يمكن تحاشي التعدد الخطي بتوسيع حجم العينة، فمثلا يمكن تحويل البيانات السنوية إلى بيانات موسمية أو شهرية إن أمكن ذلك، كما يمكن التخلص من التعدد الخطي بإسقاط (حذف) المتغير المسبب لهذا المشكل لكن هذه العملية يمكن أن تخلق مشاكل أخرى، وهناك من يقترح إدخال معلومات إضافية للنموذج.

إن وجود التعدد الخطي يجعل من الصعب فصل آثار المتغيرات المختلفة، ومنه نحتاج إلى معلومات خاصة تساعدنا على فصل أثر كل متغير لوحده، ويكون ذلك عن طريق فرض قيود على بعض المعالم بناءا على المعلومات المسبقة للنظرية الاقتصادية.

2. الارتباط الذاتي بين الأخطاء:

من بين الافتراضات الكلاسيكية التي وضعناها من قبل لتقدير معالم نموذج الانحدار، هو استقلال القيمة المقدرة لحد الخطأ في فترة زمنية معينة عن القيمة المقدرة لحد الخطأ في فترة زمنية سابقة لها. أي:

$$Cov(\varepsilon_i, \varepsilon_j) = 0, \quad \forall i \neq j$$

وإذا تم إسقاط هذا الافتراض فإن ذلك يدل على وجود ما يسمى بالارتباط الذاتي حيث أن مصفوفة التباين–التباين المشترك $E(\varepsilon\varepsilon') = \Omega_\varepsilon \neq \sigma_\varepsilon^2 I$ لا تحتوي على الصفر خارج القطر الأول و كنتيجة لذلك:

$$\Omega_{\hat{\beta}} = E((\hat{\beta}-\beta)(\hat{\beta}-\beta)') = (X'X)^{-1} X'E(\varepsilon\varepsilon')X(X'X)^{-1}$$
$$= (X'X)^{-1}(X'\Omega_\varepsilon X)(X'X)^{-1} \neq \sigma_\varepsilon^2 (X'X)^{-1}$$

يتم استعمال طريقة المربعات الصغرى المعممة GLS لتقدير شعاع المعالم β والذي ينبغي أن يكون لديه نفس الخصائص الإحصائية لأي مقدر:

$$\hat{\beta} = (X'\Omega_\varepsilon^{-1} X)^{-1}(X'\Omega_\varepsilon^{-1} Y)$$
$$\Omega_{\hat{\beta}} = (X'\Omega_\varepsilon^{-1} X)^{-1}$$

عندما تكون الفرضيات الأساسية للنموذج محققة، فإن:

$$\hat{\beta}_{GLS} = (X'\Omega_\varepsilon^{-1} X)^{-1}(X'\Omega_\varepsilon^{-1} Y) = \left(X'\left(\sigma_\varepsilon^2 I\right)^{-1} X\right)^{-1}\left(X'\left(\sigma_\varepsilon^2 I\right)^{-1} Y\right) = (X'X)^{-1}(X'Y) = \hat{\beta}_{OLS}$$

في هذه الحالة، المقدر المتحصل عليه بطريقة المربعات الصغرى المعممة هي نفسه المقدر بطريقة المربعات الصغرى العادية.

1.2. أسبابه وطرق كشفه

ينشأ الارتباط الذاتي من عدة أسباب منها:

- ❖ إهمال بعض المتغيرات التفسيرية في النموذج المراد تقديره.
- ❖ الصياغة الرياضية الخاطئة للنموذج.
- ❖ عدم دقة بيانات السلاسل الزمنية.

أما وجوده يؤثر سلبا على نتائج المربعات الصغرى العادية من حيث:

- ❖ سوف تكون المقدرات غير متحيزة.
- ❖ تباين مقدرات النموذج سوف لا يكون أقل ما يمكن.

لذلك تستعمل عدة اختبارات للكشف على هذا الاختلال منها ما يلي:

1.1.2. اختبار درابين واتسون Durbin-Watson test (1950 et 1951):

يعتبر اختبار Durbin-Watson من أهم الاختبارات الشائعة المستخدمة في اكتشاف الارتباط الذاتي من الدرجة الأولى حسب الشكل:

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t \ , \ \ u_t \sim N(0,\sigma_u^2)$$

ويهدف إلى اختبار الفرضيات التالية:

$$H_0 : \rho = 0$$
$$H_1 : \rho \neq 0$$

لاختبار فرضية العدم H_0 يجب حساب إحصائية دربين واتسون DW:

$$DW = \frac{\sum_{t=2}^{n}\left(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}\right)^2}{\sum_{t=1}^{n}\hat{\varepsilon}_{t-1}^2}$$

يمكن كتابة الإحصائية أيضا بدلالة مقدر معامل الارتباط ρ، لدينا:

$$DW = \frac{\sum_{t=1}^{n}\hat{\varepsilon}_t^2 + \sum_{t=2}^{n}\hat{\varepsilon}_{t-1}^2 - 2\sum_{t=1}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n}\hat{\varepsilon}_{t-1}^2}$$

نلاحظ أن: $\sum_{t=2}^{n}\hat{\varepsilon}_{t-1}^2 \cong \sum_{t=1}^{n}\hat{\varepsilon}_t^2$ إذن:

$$DW \cong \frac{2\sum_{t=2}^{n}\hat{\varepsilon}_{t-1}^2}{\sum_{t=1}^{n}\hat{\varepsilon}_{t-1}^2} - \frac{2\sum_{t=1}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n}\hat{\varepsilon}_{t-1}^2}$$

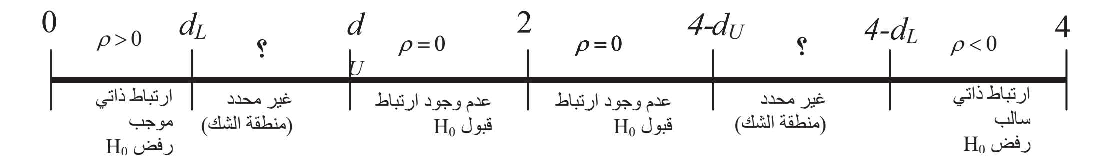
نعلم أن:
$$\hat{\rho} = \frac{\sum_{t=1}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n} \hat{\varepsilon}_{t-1}^2}$$

ومنه:
$$DW \cong 2(1-\hat{\rho})$$

حيث أن DW تمثل القيمة المحسوبة للاختبار وتأخذ قيمها بين 0 و4. ويتضح من المعادلة السابقة أنه إذا كانت $\rho = 0$ فإن $DW \cong 2$.

ويوضح الشكل التالي قيم d (القيم المجدولة للاختبار)، التي تشير إلى وجود أو عدم وجود ارتباط ذاتي من الدرجة الأولى موجب أو سالب، أو تجعل نتيجة الاختبار غير محددة، وتوجد قيم كل من الحدين الأعلى والأدنى لـ d . (d_L,d_U) في الجدول الإحصائي لتوزيع دربين واتسون.

الشكل رقم (1): مناطق القبول والرفض لاختبار Durbin-Watson



بالاعتماد على الشكل رقم (1) يمكن أن تُستخرج نتيجة اختبار DW كالتالي:

❖ إذا كانت $DW < d_L$ أو $DW > 4-d_L$ يرفض H_0.

❖ إذا كانت $4-d_U > DW > d_U$ يقبل H_0.

❖ إذا كانت $d_L \leq DW \leq d_U$ أو $4-d_U \leq DW \leq 4-d_L$ تكون نتيجة الاختبار غير محددة، ومن ثم يجب إضافة بيانات أكثر.

لا يمكن استعمال هذا الاختبار إلا في ظل الشروط التالية:

❖ يجب أن يكون النموذج متضمنا للمعلم الثابت β_0

❖ النموذج المقدر لا يتضمن متغيرات تابعة ذات فترات إبطاء كمتغيرات مستقلة[1]

1- قد يتضمن هذا النوع من النماذج متغيرات تابعة ذات فترات إبطاء كمتغيرات مستقلة، سيتم التطرق إلى هذا النوع من النماذج في الفصل الرابع.

لا يختبر دربين واتسون إلا الارتباط الذاتي من الدرجة الأولى.

2.1.2.. اختبار Breusch-Godfrey[1]

يرتكز هذا الاختبار على مضاعف لاغرانج و الذي يسمح باختبار وجود ارتباط ذاتي من درجة أكبر من الواحد. نموذج الانحدار الذاتي للأخطاء من الدرجة p يكتب على الشكل التالي:

$$\varepsilon_t = \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + \ldots + \rho_p\varepsilon_{t-p} + u_t$$

ليكن النموذج العام حيث أن الأخطاء مرتبطة ذاتيا:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \ldots + \beta_k X_{tk} + \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + \ldots + \rho_p\varepsilon_{t-p} + u_t$$

هناك ثلاث خطوات لإجراء هذا الاختبار:

- ❖ تقدير النموذج العام بطريقة المربعات الصغرى ثم حساب البواقي $\hat{\varepsilon}_t$
- ❖ تقدير المعادلة الوسيطية التالية:

$$\hat{\varepsilon}_t = \beta_0 + \beta_1 X_{t1} + \ldots + \beta_k X_{tk} + \rho_1\hat{\varepsilon}_{t-1} + \rho_2\hat{\varepsilon}_{t-2} + \ldots + \rho_p\hat{\varepsilon}_{t-p} + u_t$$

ثم حساب معامل التحديد الخاص بهذه المعادلة R^2. نذكر أن باستعمال هذه المعادلة، سنفقد p مشاهدة.

- ❖ فرضية استقلالية الأخطاء H_0 التي ينبغي اختبارها هي:

$$H_0 : \rho_1 = \rho_2 = \ldots = \rho_p = 0$$

الإحصائية $LM = (n-p)\times R^2$ تتبع توزيع χ^2 بدرجة حرية p. إذا كان $(n-p)\times R^2$ أكبر من $\chi^2(p)$ (القيمة الحرجة لتوزيع χ^2 بنسبة معنوية α)، فإننا نرفض H_0 فرضية استقلالية الأخطاء.

1- أنظر Breusch (1978) و Godfrey (1978).

2.2. طرق التقدير في حالة وجود ارتباط ذاتي بين الأخطاء:

إذا أخذنا بعين الاعتبار الارتباط الذاتي بين الأخطاء، النموذج الخطي العام يكتب على الشكل التالي:

$$Y = X\beta + \varepsilon$$

مع: $$\varepsilon_t = \rho\,\varepsilon_{t-1} + u_t,\ |\rho| < 1$$

تعبر هذه المعادلة عن سيرورة الانحدار الذاتي من الدرجة الأولى *AR(1)* الذي يحقق الفرضيات التالية:

$$E(u_t) = 0$$
$$E(u_t^2) = \sigma_u^2, \forall t$$
$$\mathrm{cov}(u_t, u_{t'}) = 0, \forall t \neq t'$$
$$\mathrm{cov}(u_t, \varepsilon_{t-1}) = 0, \forall t$$

من نموذج الانحدار الذاتي، نعوض ε_{t-1} بما يساويه، نحصل على:

$$\varepsilon_t = \rho(\rho\varepsilon_{t-2} + u_{t-1}) + u_t = \rho^2\varepsilon_{t-2} + \rho u_{t-1} + u_t$$

نعوض ε_{t-2} بما يساويها، نحصل أيضا:

$$\varepsilon_t = \rho^2(\rho\varepsilon_{t-3} + u_{t-2}) + \rho u_{t-1} + u_t = \rho^3\varepsilon_{t-3} + \rho^2 u_{t-2} + \rho u_{t-1} + u_t$$

في الأخير، لدينا العادلة:

$$\varepsilon_t = u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \rho^3 u_{t-3} + \ldots$$

تؤول هذه السيرورة إلى الصفر لأن: $|\rho| < 1$.

نقوم إذن بدراسة خصائص ε_t.

لدينا $E(\varepsilon_t) = 0$ و بتربيع الخطأ ε_t، نحصل على:

$$\varepsilon_t^2 = u_t^2 + \rho^2 u_{t-1}^2 + \rho^4 u_{t-2}^2 + \rho^6 u_{t-3}^2 + \ldots$$

ثم بإدخال التوقع الرياضي على الطرفين:

$$E(\varepsilon_t^2) = E(u_t^2) + \rho^2 E(u_{t-1}^2) + \rho^4 E(u_{t-2}^2) + \rho^6 E(u_{t-3}^2) + \ldots$$

نعلم أن: $E(u_t^2) = E(u_{t-1}^2) = E(u_{t-2}^2) = E(u_{t-3}^2) = ... = \sigma_u^2, \forall t$

ومنه: $\text{var}(\varepsilon_t) = E(\varepsilon_t^2) = (1+\rho^2+\rho^4+\rho^6+...)\sigma_u^2$

إذن: $\text{var}(\varepsilon_t) = E(\varepsilon_t^2) = \dfrac{\sigma_u^2}{1-\rho^2}$

وباستعمال خصائص u_t، نقوم بحساب التباين المشترك للخطأ ε_t، لدينا:

$$\varepsilon_t\varepsilon_{t-1} = \left(u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \rho^3 u_{t-3} + ...\right)\left(u_{t-1} + \rho u_{t-2} + \rho^2 u_{t-3} + \rho^3 u_{t-4} + ...\right)$$
$$= \rho u_{t-1}^2 + \rho^3 u_{t-2}^2 + \rho^5 u_{t-3}^2 + ...$$

بإدخال التوقع الرياضي على الطرفين، نحصل في الأخير على:

$$E(\varepsilon_t\varepsilon_{t-1}) = \frac{\rho\sigma_u^2}{1-\rho^2}$$

نفس الشيء مع: $E(\varepsilon_t\varepsilon_{t-2})$

$$E(\varepsilon_t\varepsilon_{t-2}) = \frac{\rho^2\sigma_u^2}{1-\rho^2}$$

في الحالة العامة، سيكون لدينا: $E(\varepsilon_t\varepsilon_{t-i}) = \dfrac{\rho^i\sigma_u^2}{1-\rho^2}$

مصفوفة التباين-التباين المشترك للأخطاء في حالة وجود ارتباط ذاتي من الدرجة الأولى تكتب على الشكل التالي:

$$\Omega_\varepsilon = E(\varepsilon\varepsilon') = \frac{\sigma_u^2}{1-\rho^2}\begin{pmatrix} 1 & \rho & \rho^2 & ... & \rho^{n-1} \\ \rho & 1 & \rho & ... & \rho^{n-2} \\ \rho^2 & \rho & 1 & ... & \rho^{n-3} \\ ... & ... & ... & ... & ... \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & ... & 1 \end{pmatrix}$$ مع $|\rho| < 1$

نذكر أن مقدر طريقة المربعات الصغرى يكتب على الصيغة التالية:

$$\hat{\beta} = (X'\Omega_\varepsilon^{-1}X)^{-1}(X'\Omega_\varepsilon^{-1}Y)$$

و هذا يعني أن معكوس مصفوفة التباين-التباين المشترك للأخطاء معرف كما يلي:

$$\Omega_{\varepsilon}^{-1} = \frac{1}{\sigma_u^2}\begin{pmatrix} 1 & -\rho & 0 & 0 & \dots & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \dots & \dots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & 1+\rho^2 & -\rho \\ 0 & \dots & \dots & \dots & 0 & -\rho & 1 \end{pmatrix}$$

من الملاحظ أن قيمتي ρ و σ_u^2 غير معروفتين وهذا يعني أنه ينبغي إيجاد مقدر لكل منهما وذلك بالبحث عن مصفوفة M حيث النموذج $MY = MX\beta + M\varepsilon$ يحقق جميع الفرضيات الأساسية.

لدينا:

$$E((M\varepsilon)(M\varepsilon)') = E(M\varepsilon\varepsilon' M') = ME(\varepsilon\varepsilon')M' = M\Omega_{\varepsilon}M' = \sigma_{\varepsilon}^2 I$$

إذن نحدد المقدر BLUE كما يلي: $\hat{\beta} = \left(X'M'MX\right)^{-1} X'M'MY$

و بالنتيجة التالية: $M'M = \lambda\Omega_{\varepsilon}^{-1} = \sigma_u^2\Omega_{\varepsilon}^{-1} \quad \forall\lambda$

نحصل على المصفوفة:

$$\underset{(n-1,n)}{M} = \begin{pmatrix} -\rho & 1 & 0 & 0 & \dots & 0 \\ 0 & -\rho & 1 & 0 & \dots & 0 \\ 0 & 0 & -\rho & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & -\rho & 1 \end{pmatrix}$$

التي تعتبر كاستنتاج للمصفوفة:

$$M'M = \begin{pmatrix} \rho^2 & -\rho & 0 & 0 & \dots & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \dots & \dots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & 1+\rho^2 & -\rho \\ 0 & \dots & \dots & \dots & 0 & -\rho & 1 \end{pmatrix}$$

و هي نفسها المصفوفة $\sigma_u^2\Omega_\varepsilon^{-1}$. وعليه تؤول طريقة المربعات الصغرى المعممة إلى طريقة المربعات الصغرى العادية، حيث يتم تقدير النموذج العام المصحح من الارتباط الذاتي عن طريق تحويل المتغيرات عن طريق شبه الفروقات من الدرجة الأولى، لدينا:

$$MY = \begin{pmatrix} Y_2 - \rho Y_1 \\ Y_3 - \rho Y_2 \\ \ldots \\ Y_n - \rho Y_{n-1} \end{pmatrix} \quad \text{و} \quad MX_j = \begin{pmatrix} X_{2,j} - \rho X_{1,j} \\ X_{3,j} - \rho X_{2,j} \\ \ldots \\ X_{n,j} - \rho X_{n-1,j} \end{pmatrix}$$

حيث: $j = 1,2,\ldots,k$

عند استخدام شبه الفروقات، نفقد المشاهدة الأولى لكل متغير ولتجنب ضياعها، نضع:

$$Y_1^* = Y_1\sqrt{1-\rho^2}$$
$$X_{1,j}^* = X_{1,j}\sqrt{1-\rho^2}$$

يكتب النموذج المصحح على النحو التالي:

$$Y_t - \rho Y_{t-1} = \beta_0(1-\rho) + \beta_1(X_{t1} - \rho X_{t-1,1}) + \beta_2(X_{t2} - \rho X_{t-1,2}) + \ldots + \beta_k(X_{tk} - \rho X_{t-1,k}) + \varepsilon_t - \rho\varepsilon_{t-1}$$

كما ذكرنا، يتم تقدير النموذج الأخير بطريقة المربعات الصغرى العادية. المشكل الأساسي الذي نواجهه يتمثل في تقدير معامل الارتباط الذاتي بين الأخطاء من الدرجة الأولى. هناك إذن عدة طرق للتقدير منها ما يلي:

1.2.2. تقدير ρ عن طريق إحصائية Durbin-Watson:

الخطوة الأولى: تقدير ρ انطلاقا من إحصائية DW، حيث: $\hat{\rho} \cong 1 - DW/2$

الخطوة الثانية: تقدير النموذج التالي بعد إجراء التعديلات على المشاهدات بحساب شبه الفروقات:

$$Y_t - \hat{\rho}Y_{t-1} = \beta_0(1-\hat{\rho}) + \beta_1(X_{t1} - \hat{\rho}X_{t-1,1}) + \beta_2(X_{t2} - \hat{\rho}X_{t-1,2}) + \ldots + \beta_k(X_{tk} - \hat{\rho}X_{t-1,k}) + u_t$$

أي: $Y_t^* = \beta_0^* + \beta_1 X_{t1}^* + \beta_2 X_{t2}^* + ... + \beta_k X_{tk}^* + u_t$

المعالم المقدرة بطريقة المربعات الصغرى هي: $\hat{\beta}_1,...,\hat{\beta}_k$ و $\hat{\beta}_0^* = \hat{\beta}_0(1-\hat{\rho})$.

2.2.2. تقدير ρ بطريقة Theil-Nagar:

اقترح Theil و Nagar تقديرا لـ ρ . من خلال العلاقة التالية:

$$\hat{\rho} = \frac{n^2\left[1-\left(DW/2\right)\right]+\left(k+1\right)^2}{n^2-\left(k+1\right)^2}$$

حيث k هي عدد المتغيرات المستقلة. نستخدم نفس الخطوات لتقدير معالم النموذج بطريقة المربعات الصغرى.

3.2.2. طريقة Cochrane-Orcutt

اقترح Cochrane و Orcutt تقديرا بإعطاء قيمة ابتدائية لـ ρ . بواسطة القيم المقدرة لحد الخطأ.

الخطوة الأولى: إعطاء قيمة ابتدائية لمعامل الارتباط وذلك بتقنية تقدير مباشرة: $\hat{\rho} = \frac{\sum \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sum \hat{\varepsilon}_{t-1}^2}$ ، فليكن:

$$\hat{\rho}_0 = \hat{\rho}$$

الخطوة الثانية: تقدير النموذج التالي بطريقة المربعات الصغرى العادية:

$$Y_t - \hat{\rho}_0 Y_{t-1} = \beta_0(1-\hat{\rho}_0) + \beta_1(X_{t1} - \hat{\rho}_0 X_{t-1,1}) + \beta_2(X_{t2} - \hat{\rho}_0 X_{t-1,2}) + ... + \beta_k(X_{tk} - \hat{\rho}_0 X_{t-1,k}) + u_t$$

المعالم المقدرة هي: $\hat{\beta}_1,...,\hat{\beta}_k$ و $\hat{\beta}_0^* = \hat{\beta}_0(1-\hat{\rho})$.

الخطوة الثالثة: إعادة تقدير ρ ببواقي تقدير جديدة $\hat{\varepsilon}_t^1$ حيث:

$$\hat{\varepsilon}_t^{(1)} = Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_{t1} - ... - \hat{\beta}_k X_{tk}$$

$$\hat{\rho}_1 = \frac{\sum \hat{\varepsilon}_t^{(1)} \hat{\varepsilon}_{t-1}^{(1)}}{\sum (\hat{\varepsilon}_t^{(1)})^2}$$

الخطوة الرابعة: تقدير النموذج التالي على المتغيرات ذات شبه الفروقات:

$$Y_t-\hat{\rho}_1 Y_{t-1}=\beta_0(1-\hat{\rho}_1)+\beta_1(X_{t1}-\hat{\rho}_1 X_{t-1,1})+\beta_2(X_{t2}-\hat{\rho}_1 X_{t-1,2})+...+\beta_k(X_{tk}-\hat{\rho}_1 X_{t-1,k})+u_t$$

ثم نعيد تقدير ρ مرة أخرى ببواقي تقدير جديدة $\hat{\varepsilon}_t^{(2)}$ فنحصل على تقدير لـ $\hat{\rho}_2$.
نكرر العملية مرات أخرى إلى غاية سكون مقدرات النموذج (عادة نكرر العملية ثلاث أو أربع مرات).

4.2.2. طريقة Hildreth-Lu:

الخطوة الأولى: تحديد نوع الارتباط (موجب أو سالب) بواسطة إحصائية Durbin-Watson.

الخطوة الثانية: نحدد مجالا للقيم الممكنة لمعامل الارتباط ρ . نختار قيما تنتمي إلى المجال $[0,1]$ إذا كان المعامل موجبا و قيما تنتمي إلى هذا المجال $[-1,0]$ إذا كان سالبا.على سبيل المثال يكون لدينا $\rho=\{0.1;0.2;...;0.9;1\}$ إذا اعتبرنا أن معامل الارتباط موجب فيتم اختيار درجة سلم 0.1 أو 0.01 ومع كل قيمة يتم تقدير النموذج:

$$Y_t-\hat{\rho}_i Y_{t-1}=\beta_0(1-\hat{\rho}_i)+\beta_1(X_{t1}-\hat{\rho}_i X_{t-1,1})+\beta_2(X_{t2}-\hat{\rho}_i X_{t-1,2})+...+\beta_k(X_{tk}-\hat{\rho}_i X_{t-1,k})+u_t$$

ونأخذ قيمة ρ القيمة التي عندها يكون مجموع مربعات البواقي $\sum_t \hat{u}_t^2$ أصغر ما يمكن.

مثال 2:

نقوم بدراسة العلاقة بين معدل الفائدة طويل الأجل Y_i بدلالة معدل الفائدة قصير الأجل X_i . الجدول التالي يحتوي على معطيات شهرية لمعدلات الفائدة في الولايات المتحدة الأمريكية.

الجدول (2): معدلات الفائدة قصيرة و طويلة الأجل

X_i	Y_i	السنة	X_i	Y_i	السنة
5.61	7.05	2005:4	7.77	6.56	2004:1
5.23	7.01	2005:5	7.12	6.54	2004:2
5.34	6.86	2005:6	7.96	6.81	2004:3
6.13	6.89	2005:7	8.33	7.04	2004:4
6.44	7.11	2005:8	8.23	7.1	2004:5
6.42	7.28	2005:9	7.90	7.02	2004:6
5.96	7.29	2005:10	7.55	7.18	2004:7
5.48	7.21	2005:11	8.96	7.33	2004:8
5.44	7.17	2005:12	8.06	7.30	2004:9
4.87	6.93	2006:1	7.46	7.22	2004:10
4.88	6.92	2006:2	7.47	6.93	2004:11
5.00	6.88	2006:3	7.15	6.77	2004:12
4.86	6.73	2006:4	6.26	6.68	2005:1
5.20	7.01	2006:5	5.50	6.66	2005:2
5.41	6.92	2006:6	5.49	6.77	2005:3

المطلوب:

- تقدير انحدار معدل الفائدة طويل الأجل Y_i على معدل الفائدة قصير الأجل ثم اختبار استقلالية الأخطاء معتمدا على مختلف الإحصائيات المذكورة سابقا.
- تصحيح النموذج من الارتباط الذاتي إن وجد.

الخطوة الأولى: تقدير النموذج بطريقة المربعات الصغرى، نحصل على النتائج باستعمال البرنامج RATS 5.04:

```
Linear Regression - Estimation by Least Squares
Dependent Variable Y
Usable Observations     30     Degrees of Freedom    28
Centered R**2      0.045827      R Bar **2   0.011750
Uncentered R**2    0.999068      T x R**2      29.972
Mean of Dependent Variable      6.9723333333
Std Error of Dependent Variable 0.2217280617
Standard Error of Estimate      0.2204215930
Sum of Squared Residuals        1.3603990021
Regression F(1,28)                    1.3448
Significance Level of F           0.25598417
Durbin-Watson Statistic             0.446190

   Variable                    Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                6.7286242502 0.2139755259     31.44577  0.00000000
2.  X                       0.0377882597 0.0325858628      1.15965  0.25598417
```

<u>الخطوة الثانية</u>: اختبار استقلالية الأخطاء

نلاحظ أن إحصائية دربين–واتسون التي تساوي 0.4461 تقع في منطقة رفض H_0 أي $0 < DW < d_L$ حيث $d_L = 1.35; d_U = 1.49$ بنسبة معنوية 0.05، $n = 30$ و $k = 1$. إذن هناك ارتباط ذاتي موجب بين الأخطاء. يمكن التأكد من ذلك باستعمال إحصائية Breusch-Godfrey. لدينا المعادلة الوسيطية المقدرة:

$$\hat{\varepsilon}_i = -0.07 + 0.014X_i + 0.93\hat{\varepsilon}_{t-1} - 0.35\hat{\varepsilon}_{t-2}$$
$$(5.14) \quad (-2.08)$$

$$R^2 = 0.6077$$

القيم التي بين قوسين هي قيم ستيودنت. تحت ظل قبول فرضية استقلالية الأخطاء H_0، معاملا $\hat{\varepsilon}_{t-1}, \hat{\varepsilon}_{t-2}$ يساويان معنويا الصفر. نلاحظ أن إحصائية مضاعف لاغرانج:

$$LM = (n-2) \times R^2 = (30-2) \times 0.6077 = 17.0168$$

أكبر تماما من القيمة الحرجة لتوزيع χ^2 بدرجة حرية 2 و نسبة معنوية 0.05 ($\chi^2_{0.05}(2) = 5.99$)، أي نرفض الفرضية H_0 و هذا يعني أن الأخطاء مرتبطة ذاتيا.

<u>الخطوة الثالثة</u>: تصحيح النموذج في حالة الارتباط الذاتي بين الأخطاء

الطريقة الأولى: نستعمل طريقة دربين-واتسون المباشرة حيث نقوم أولا بحساب معامل الارتباط الذاتي $\hat{\rho} \approx 1 - \frac{DW}{2} = 1 - 0.4461/2 = 0.7769$ ثم نحول المتغيرات الأصلية إلى متغيرات ذات شبه الفروقات:

$$Y_t^* = Y_t - \hat{\rho} Y_{t-1}$$
$$X_t^* = X_t - \hat{\rho} X_{t-1}$$

أما المشاهدة الأولى لكل منهما يتم حسابها كما يلي:

$$Y_1^* = Y_1 \sqrt{1 - \hat{\rho}^2}$$
$$X_1^* = X_1 \sqrt{1 - \hat{\rho}^2}$$

البيانات المحولة تظهر في الجدول أدناه:

الجدول (3): المعطيات الإحصائية المحولة

$Y_t^* = Y_t - \hat{\rho} Y_{t-1}$	$X_t^* = X_t - \hat{\rho} X_{t-1}$	Y_t	X_t	السنة
2.60	4.89	6.56	7.77	2004:1
1.44	1.08	6.54	7.12	2004:2
1.72	2.42	6.81	7.96	2004:3
1.74	2.14	7.04	8.33	2004:4
1.63	1.75	7.1	8.23	2004:5
1.50	1.50	7.02	7.90	2004:6
1.72	1.41	7.18	7.55	2004:7
1.75	3.09	7.33	8.96	2004:8
1.60	1.09	7.30	8.06	2004:9
1.54	1.19	7.22	7.46	2004:10
1.32	1.67	6.93	7.47	2004:11
1.38	1.34	6.77	7.15	2004:12
1.42	0.70	6.68	6.26	2005:1
1.47	0.63	6.66	5.50	2005:2
1.59	1.21	6.77	5.49	2005:3
1.79	1.34	7.05	5.61	2005:4
1.53	0.87	7.01	5.23	2005:5
1.41	1.27	6.86	5.34	2005:6
1.56	1.98	6.89	6.13	2005:7
1.75	1.67	7.11	6.44	2005:8
1.75	1.41	7.28	6.42	2005:9

$Y_t^* = Y_t - \hat{\rho}Y_{t-1}$	$X_t^* = X_t - \hat{\rho}X_{t-1}$	Y_t	X_t	السنة
1.63	0.97	7.29	5.96	2005:10
1.54	0.84	7.21	5.48	2005:11
1.56	1.18	7.17	5.44	2005:12
1.35	0.64	6.93	4.87	2006:1
1.53	1.09	6.92	4.88	2006:2
1.50	1.20	6.88	5.00	2006:3
1.38	0.97	6.73	4.86	2006:4
1.78	1.42	7.01	5.20	2006:5
1.47	1.37	6.92	5.41	2006:6

نقوم الآن بتقدير النموذج الجديد المحول $Y_t^* = \beta_0^* + \beta_1 X_t^* + u_t$ بطريقة المربعات الصغرى حيث $\beta_0^* = \beta_0(1-\hat{\rho})$ و $u_t = \varepsilon_t - \hat{\rho}\varepsilon_{t-1}$، نحصل على النتائج التالية:

```
Linear Regression - Estimation by Least Squares
Dependent Variable DY
Usable Observations     30     Degrees of Freedom    28
Centered R**2      0.669552      R Bar **2    0.657751
Uncentered R**2    0.993275      T x R**2       29.798
Mean of Dependent Variable      1.5983333333
Std Error of Dependent Variable 0.2343013582
Standard Error of Estimate      0.1370711912
Sum of Squared Residuals        0.5260783205
Regression F(1,28)                   56.7335
Significance Level of F           0.00000003
Durbin-Watson Statistic             1.755006

   Variable                       Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                   1.2586509473 0.0515759131     24.40385  0.00000000
2.  DX                         0.2298775453 0.0305194494      7.53217  0.00000003
```

حيث وضعنا $DY = Y^=$ و $DX = X^*$. الثابتة المقدرة تحسب كما يلي:

$$\hat{\beta}_0^* = \hat{\beta}_0(1-\hat{\rho}) \Rightarrow \hat{\beta}_0 = \frac{\hat{\beta}_0^*}{(1-\hat{\rho})} = \frac{1.2584}{1-0.7769} = 5.64$$

نلاحظ أن النموذج المصحح من الارتباط الداتي مقبول إحصائيا حيث أن له قدرة تفسيرية عالية وللمعالم معنوية إحصائية. يمكن القول أن إحصائية دربين-واتسون تشير إلى استقلالية تامة بين الأخطاء حيث تقع في منطقة قبول الفرضية H_0 و بالتالي تم التخلص من الارتباط الذاتي وذلك بتصحيح النموذج وفق طريقة دربين-واتسون.

الطريقة الثانية: طريقتا Hidreth-Lu و Cochrane-Orcutt.

في هذه الحالة يتم تقدير شعاع المعالم β و ρ بالطريقة التكرارية. النتائج معطاة أيضا باستعمال نفس برنامج الكمبيوتر:

1. التقدير بطريقة Hidreth-Lu:

```
Regression with AR1 - Estimation by Hildreth-Lu Search
Dependent Variable Y
Usable Observations     29     Degrees of Freedom    26
Centered R**2      0.674989      R Bar **2   0.649989
Uncentered R**2    0.999713      T x R**2      28.992
Mean of Dependent Variable      6.9865517241
Std Error of Dependent Variable 0.2112763232
Standard Error of Estimate      0.1249947867
Sum of Squared Residuals        0.4062161140
Durbin-Watson Statistic             1.624322

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                 6.2130001242 0.2819342680     22.03705  0.00000000
2.  X                        0.1319715005 0.0431488699      3.05852  0.00510365
*******************************************************************************
3.  RHO                      0.7555583659 0.0987214671      7.65344  0.00000004
```

2. التقدير بطريقة Cochrane-Orcutt:

```
Regression with AR1 - Estimation by Cochrane-Orcutt
Dependent Variable Y
Usable Observations     29     Degrees of Freedom    26
Centered R**2      0.674989      R Bar **2   0.649989
Uncentered R**2    0.999713      T x R**2      28.992
Mean of Dependent Variable      6.9865517241
Std Error of Dependent Variable 0.2112763232
Standard Error of Estimate      0.1249947867
Sum of Squared Residuals        0.4062161140
Durbin-Watson Statistic             1.624321

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                 6.2130007490 0.2819340161     22.03707  0.00000000
2.  X                        0.1319713411 0.0431488383      3.05851  0.00510367
*******************************************************************************
3.  RHO                      0.7555574469 0.0987215179      7.65342  0.00000004
```

يتبين من خلال هذه النتائج أن طريقتي Hidreth-Lu و Cochrane-Orcutt تعطيان تقريبا نفس النتائج إلى حد بعيد. و في هذه الحالة أيضا تخلصنا من مشكل الارتباط الذاتي بين الأخطاء حيث إحصائية دربين–واتسون تشير إلى استقلالية تامة بين بواقي التقدير. تجدر الإشارة في الأخير إلى أن هاتين الطريقتين المبنيتين أساسا على الأسلوب التكراري تعطيان أحسن النتائج بالمقارنة مع الطريقة المباشرة لدربين–واتسون.

3. عدم تجانس تباين الأخطاء Heteroscedasticity:

1.3. طبيعة عدم ثبات تباين الأخطاء، أسبابه وآثاره:

إذا كانت فرضية تجانس التباين غير محققة، فإن مصفوفة التباين–التباين المشترك للأخطاء تعرف كما يلي:

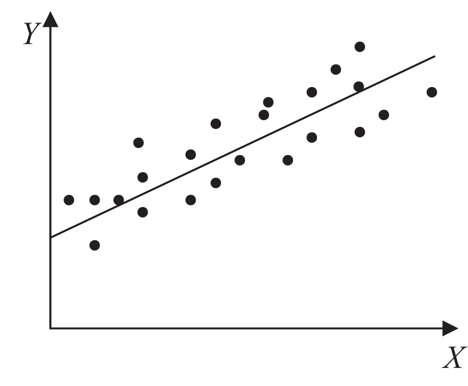
$$\Omega_\varepsilon = E(\varepsilon\varepsilon') = \begin{pmatrix} \sigma^2_{\varepsilon,1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\varepsilon,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2_{\varepsilon,n} \end{pmatrix} \neq \sigma^2_\varepsilon I_n$$

من الملاحظ أن تباينات الأخطاء ليست ثابتة على القطر الأول وبالتالي تباين الأخطاء مرتبط بقيم المتغير المستقل كما يظهر الشكل (3-3).

يوضح الشكل رقم (2) العلاقة المتوقعة بين المتغيرين التابع Y والمستقل X في حالة ثبات تباين الخطأ، ويلاحظ من خلال هذا الشكل أن تباين حد الخطأ لا يعتمد على قيم X.

الشكل رقم (2)

ثبات تباين الخطأ في نموذج الانحدار البسيط



الشكل رقم (3)

عدم ثبات تباين الخطأ في نموذج الانحدار البسيط

ويوضح الشكل رقم (3) حالة عدم ثبات التباين لحد الخطأ $E(\varepsilon_i^2) \neq \sigma^2, \forall i$ ، حيث نلاحظ أن زيادة X سوف تؤدي إلى زيادة تباين حد الخطأ، ويرتبط هذا المشكل ببيانات المقطع المستعرض Cross-section date أكثر من بيانات السلسلة الزمنية Cross-series date، حيث أن الأولى عبارة عن بيانات يتم تجميعها عن متغير ما في لحظة زمنية معينة (مثال: بيانات الإنفاق الاستهلاكي عند مستويات مختلفة لدخول الأفراد لسنة 2005)، أما بيانات السلسلة الزمنية فيتم تجميعها عن متغير ما عبر فترة زمنية معينة. وهناك عدة أسباب لعدم تجانس تباين حد الخطأ منها تحسن أساليب تجميع البيانات، وهذا يُقلل من الأخطاء المرتكبة في القياس، ومن ثم سوف يقل تباين حد الخطأ.

ويترتب على مشكلة عدم ثبات التباين عددا من الآثار تتمثل في[1]:

1. تبقى المعالم المقدرة باستخدام المربعات الصغرى متصفة بعدم التحيز والاتساق، ولكنها تفقد صفة الكفاءة.
2. تصبح التباينات المقدرة وكذلك التباينات المشتركة Covariances الخاصة بالمعالم المقدرة متحيزة وغير متسقة، ولذا فإن اختبارات الفرضيات لا تصبح دقيقة أو ملائمة.
3. بالرغم من أن التنبؤات القائمة على أساس المعالم المقدرة باستخدام المربعات الصغرى العادية تظل غير متحيزة، إلا أنها تفقد صفة الكفاءة، وهو ما يعني أنها تكون أقل مصداقية من التنبؤات الأخرى.

في حالة عدم تجانس تباين الأخطاء، مقدر BLUE بطريقة المربعات الصغرى المعممة يكتب كما يلي:

$$\hat{\beta} = (X'\Omega_\varepsilon^{-1}X)^{-1}(X'\Omega_\varepsilon^{-1}Y)$$
$$\Omega_{\hat{\beta}} = (X'\Omega_\varepsilon^{-1}X)^{-1}$$

عكس تصحيح النموذج من الارتباط الذاتي، لا توجد منهجية موحدة للتصحيح من عدم ثبات تباين الأخطاء، فالطرق مختلفة مرتبطة بسبب وجود هذا المشكل.

1- عبد القادر محمد عبد القادر عطية، ص 439.

2.3. اختبارات اكتشاف عدم تباين الخطأ:

يتم اكتشاف عدم ثبات تباين الأخطاء بواسطة عدة اختبارات منها ما يلي:

1.2.3. اختبار Goldfeld-Quandt:

بافتراض النموذج التالي $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, i = 1,...,n$، يمكن تبيان كيفية استخدام اختبار Goldfeld-Quandt في اكتشاف عدم ثبات تباين الخطأ من خلال الخطوات التالية:

- ترتيب مشاهدات X ترتيبا تصاعديا.
- استبعاد المشاهدات الوسطى لكل من X و Y،ثم تكوين مجموعتين من المشاهدات بحيث يكون لكل مجموعة على حدا معادلة خاصة بها كما يلي:

1. المجموعة الأولى: وتتمثل في المشاهدات الخاصة بكل من X و Y الواردة قبل المشاهدات التي تم استبعادها، والمعادلة الخاصة بهذه المجموعة هي:

$$Y_{1i} = a + bX_{1i} + \varepsilon_{1i}$$

2. المجموعة الثانية: وتتمثل في المشاهدات الخاصة بكل من X و Y الواردة بعد المشاهدات التي تم استبعادها، والمعادلة الخاصة بهذه المجموعة هي:

$$Y_{2i} = c + dX_{2i} + \varepsilon_{2i}$$

- تقدير معاملات المعادلتين السابقتين باستعمال المربعات الصغرى:

$$\hat{Y}_{1i} = \hat{a} + \hat{b}X_{1i}$$

$$Y_{2i} = \hat{c} + \hat{d}X_{2i}$$

- الحصول على القيم المقدرة لحد الخطأ:

$$\hat{\varepsilon}_{1i} = Y_{1i} - \hat{Y}_{1i}$$

$$\hat{\varepsilon}_{2i} = Y_{2i} - \hat{Y}_{2i}$$

- إيجاد القيمة المحسوبة لإحصائية F كما يلي:

$$F = \frac{\sum \hat{\varepsilon}_{2i}^2}{\sum \hat{\varepsilon}_{1i}^2}$$

- إيجاد درجات الحرية: $DF = \dfrac{n - m - 2(k+1)}{2}$

حيث k: عدد المتغيرات المستقلة، m: عدد المشاهدات المستبعدة.

- ❖ إيجاد القيمة المجدولة لإحصائية F عند درجات الحرية لكل من البسط والمقام، ومستوى معنوية معين.
- ❖ مقارنة بين القيم المحسوبة لإحصائية F والقيمة المجدولة لها:
  - فإذا كانت F المحسوبة أكبر من F المجدولة، نقبل الفرضية البديلة أي فرضية عدم ثبات تباين الأخطاء.
  - أما إذا كانت F المحسوبة أقل من F المجدولة، يتم قبول فرضية العدم.

لاحظ أن اختبار Goldfeld-Quandt لا يمكن تطبيقه إلا في حالة ما إذا كانت إحدى المتغيرات المستقلة هي المسببة في وجود مشكلة عدم ثبات تباين حد الخطأ.

2.2.3. اختبار White:

اقترح White (1980) اختبارا يعتمد على العلاقة بين مربعات البواقي و جميع المتغيرات المستقلة و كذا مربعاتها. يمكن إبراز خطوات هدا الاختبار كما يلي:

- ❖ تقدير النموذج العام $Y = X\beta + \varepsilon$ بطريقة المربعات الصغرى العادية ثم حساب مربعات البواقي $\hat{\varepsilon}_t^2$.
- ❖ تقدير المعادلة الوسيطية التالية:

$$\hat{\varepsilon}_t^2 = \beta_0 + \beta_1 X_{t1} + \alpha_1 X_{t1}^2 + ... + \beta_k X_{tk} + \alpha_k X_{tk}^2 + u_t$$

ثم حساب معامل التحديد الخاص بهذه المعادلة R^2.

- ❖ فرضية ثبات تباين الأخطاء H_0 التي ينبغي اختبارها هي:

$$H_0 : \beta_0 = \alpha_1 = \beta_1 = ... = \alpha_k = \beta_k = 0$$

إحصائية مضاعف لاغرانج $LM = n \times R^2$ تتبع توزيع χ^2 بدرجة حرية $2k$. إذا كان $n \times R^2$ أكبر من $\chi^2(2k)$ (القيمة الحرجة لتوزيع χ^2 بنسبة معنوية α)، فإننا نرفض H_0 أي إذا كان هناك على الأقل معامل واحد من معاملات المعادلة الوسيطية يختلف معنويا عن الصفر فإن تباين الأخطاء غير متجانس.

3.2.3 اختبار ثبات التباين الشرطي للأخطاء ARCH-LM:

تسمح نماذج ARCH[1] بنمذجة المتغيرات المالية التي تحتوي على تباين شرطي غير ثابت للأخطاء العشوائية حيث أن التطاير الشرطي الذي يعبر في الغالب عن المخاطرة غير ثابت. يعتمد إذن هدا الاختبار على مضاعف لاغرانج LM. خطوات الاختبار كالتالي:

- ❖ تقدير النموذج العام $Y = X\beta + \varepsilon$ بطريقة المربعات الصغرى العادية ثم حساب مربعات البواقي $\hat{\varepsilon}_t^2$.
- ❖ تقدير المعادلة التالية:

$$\hat{\varepsilon}_t^2 = \theta_0 + \theta_1 \hat{\varepsilon}_{t-1}^2 + \ldots + \theta_q \hat{\varepsilon}_{t-q}^2 + u_t$$

مع حساب معامل التحديد الخاص بهذه المعادلة R^2. نفقد في هذه الحالة q مشاهدة.

- ❖ فرضية ثبات التباين الشرطي للأخطاء H_0 التي ينبغي اختبارها هي:

$$H_0 : \theta_0 = \theta_1 = \ldots = \theta_q = 0$$

إحصائية مضاعف لاغرانج $LM = (n-q) \times R^2$ تتبع توزيع χ^2 بدرجة حرية q.
إذا كان $(n-q) \times R^2$ أكبر من $\chi^2(q)$ (القيمة الحرجة لتوزيع χ^2 بنسبة معنوية α)، فإننا نرفض H_0 أي إذا كان هناك على الأقل معامل واحد من معاملات معادلة ARCH يختلف معنويا عن الصفر فإن التباين الشرطي للأخطاء غير متجانس.

3.3. معالجة عدم ثبات تباين حد الخطأ:

من أبرز الطرق المستخدمة لتصحيح المشكلة هي طريقة المربعات الصغرى المرجحة، وتقوم هذه الفكرة على إعطاء القيم ذات الانحراف الأقل على خط الانحدار وزنا أكبر من القيم ذات الانحراف الأكبر في تقدير العلاقة محل الاعتبار[2]. ويتوقف شكل النموذج الأصلي المُحوَّل على نمط عدم ثبات التباين المكتشف في النموذج الأصلي المقدر.

1- Engle (1982)

2- عبد القادر محمد عبد القادر عطية، 1990، ص 452.

وبفرض أن النموذج الأصلي كان كما يلي: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i , i = 1,...,n$ فإن هناك عدة أنماط (افتراضات) لعدم ثبات تباين الأخطاء، ويختلف النموذج أو المعادلة المحولة من افتراض إلى آخر.

❖ الافتراض الأول: $E(\varepsilon_i^2) = \sigma^2 X_i^2$ وطبقا لهذا الافتراض يتم تحويل النموذج الأصلي إلى الشكل التالي:

$$\frac{Y_i}{X_i} = \frac{\beta_0}{X_i} + \beta_1 + \frac{\varepsilon_i}{X_i} = \beta_0 \frac{1}{X_i} + \beta_1 + \theta_i$$

حيث: θ_i عبارة عن حد الخطأ المحول $\frac{\varepsilon_i}{X_i}$

وبإجراء انحدار $\frac{Y_i}{X_i}$ على $\frac{1}{X_i}$ مستخدما طريقة المربعات الصغرى العادية نحصل على:

$$\left(\frac{\hat{Y}_i}{X_i}\right) = \hat{\beta}_0 \frac{1}{X_i} + \hat{\beta}_1$$

وبضرب المعادلة المحولة المقدرة السابقة في X_i يتم الحصول على النموذج الأصلي $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ بعد معالجة عدم ثبات التباين σ_ε^2، ويتضح مما سبق أن الحد الثابت في النموذج المحول (β_1) هو عبارة عن ميل معامل الانحدار للنموذج الأصلي، وميل معامل الانحدار للنموذج المحول هو عبارة عن الحد الثابت في النموذج الأصلي.

❖ الافتراض الثاني: $E(\varepsilon_i^2) = \sigma^2 X_i$ وطبقا لهذا الافتراض يتم تحويل النموذج الأصلي إلى المعادلة التالية:

$$\frac{Y_i}{\sqrt{X_i}} = \frac{\beta_0}{\sqrt{X_i}} + \beta_1 \sqrt{X_i} + \frac{\varepsilon_i}{\sqrt{X_i}} = \beta_0 \frac{1}{\sqrt{X_i}} + \beta_1 \sqrt{X_i} + \omega_i$$

حيث ω_i عبارة عن حد الخطأ المحول $\frac{\varepsilon_i}{\sqrt{X_i}}$، $X_i > 0$

وبنفس الحالة الأولى نُجري انحدار $\frac{Y_i}{\sqrt{X_i}}$ على $\frac{1}{\sqrt{X_i}}$، $\sqrt{X_i}$ بواسطة المربعات الصغرى العادية.

❖ الافتراض الثالث: $E(\varepsilon_i^2) = \sigma^2 \ Y_i^2$ ، وطبقا لهذا الافتراض تكون المعادلة المحولة من الشكل:

$$\frac{Y_i}{Y_i} = \frac{\beta_0}{Y_i} + \beta_1 \frac{X_i}{Y_i} + \frac{\varepsilon_i}{Y_i} = \beta_0 \frac{1}{Y_i} + \beta_1 \frac{X_i}{Y_i} + \varphi$$

❖ الافتراض الرابع: $E(\varepsilon_i^2) = \sigma^2 \ |\hat{\varepsilon}_i|$ ، ويتضمن هذا الافتراض أن تباين حد الخطأ دالة خطية لبواقي طريقة المربعات الصغرى العادية، وطبقا لهذا تكون المعادلة المقدرة كما يلي:

$$\frac{Y_i}{\sqrt{|\hat{\varepsilon}_i|}} = \beta_0 \frac{1}{\sqrt{|\hat{\varepsilon}_i|}} + \beta_1 \frac{X_i}{\sqrt{|\hat{\varepsilon}_i|}} + \frac{\varepsilon_i}{\sqrt{|\hat{\varepsilon}_i|}}$$

❖ الافتراض الخامس: التحويلات اللوغاريتمية، إن تحويل النموذج الأصلي إلى الصيغة اللوغاريتمية المزدوجة سوف يؤدي غالبا إلى تقليل درجة عدم ثبات تباين حد الخطأ، ومن ثم طبقا لهذا الافتراض تكون المعادلة المحولة المناسبة للنموذج الأصلي كما يلي: $LnY_i = \beta_0 + \beta_1 LnX_i + \varepsilon_i$.

مثال 3:

إذا كان لدينا 31 مشاهدة تتعلق بالدخل X_i و الادخار Y_i و المطلوب اختبار تجانس تباين الأخطاء للنموذج $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$. الجدول التالي يعطي تطور كل من الدخل و الادخار خلال 31 سنة.

الجدول (4): المعطيات الإحصائية حول الادخار و الدخل خلال 31 سنة

Y_t	X_t	i
264	8777	1
90	9954	2
122	10979	3
588	15522	4
779	18575	5
1222	21163	6
1654	25604	7
1400	26500	8

Y_t	X_t	i
1829	27670	9
2200	28300	10
2017	27430	11
2105	29560	12
105	9210	13
107	11912	14
503	13499	15
898	16730	16
819	19635	17
1702	22880	18
1600	28150	19
2450	32500	20
2570	35250	21
1720	33500	22
1900	36000	23
131	10508	24
406	12747	25
431	14269	26
951	17663	27
1578	24127	28
2250	32100	29
2100	36200	30
2300	38200	31

نستعمل أولا اختبار Goldfeld-Quandt لهذا الغرض. نقوم أولا بتطبيق طريقة المربعات الصغرى لتقدير دالة الادخار الكلي، نحصل على النتائج التالية:

$$\hat{Y}_i = -649.71 + 0.08X_i$$

$$(-5.46) \quad (17.26)$$

$$R^2 = 0.911; RSS = 1798234.34; DW = 1.49$$

القيم التي بين قوسين هي قيم ستيودنت. لتطبيق هذه الإحصائية نتبع الخطوات التالية:

الخطوة الأولى: يتم ترتيب المشاهدات ترتيبا تصاعديا حسب قيم الدخل X_i.

الخطوة الثانية: بما أن عدد المشاهدات هو 31، فإنه يتم اختيار عدد المشاهدات المستبعدة $m=9$، و هي تقع بين القيمتين $\frac{11}{3}$ و $\frac{n}{4}$ لكي يكون عدد المشاهدات المتبقية زوجيا، حيث تحذف المشاهدات التسع من الجزء الأوسط من المشاهدات بعد ترتيبها و عليه نحصل على مجموعتين كل منهما يساوي 11 أي $11 = \frac{n-m}{2}$.

الخطوة الثالثة: نقوم بتقدير النموذج على الجزء الأول من المشاهدات بعد ترتيبها تصاعديا و استبعاد المجموعة m، لدينا:

$$\hat{Y}_i = -738.84 + 0.08X_i$$
$$R^2 = 0.78; RSS_1 = 144771.5; n = 11$$

الخطوة الرابعة: نقوم بتقدير النموذج على الجزء الثاني من المشاهدات بعد ترتيبها تصاعديا و استبعاد المجموعة m، لدينا:

$$\hat{Y}_i = -1050.79 + 0.032X_i$$
$$R^2 = 0.16; RSS_2 = 769899.3; n = 11$$

الخطوة الخامسة: لاتخاذ القرار، يكفي حساب إحصائية فيشر، حيث:

$$F_c = \frac{RSS_2}{RSS_1} = \frac{769899.3}{144771.5} = 5.31$$

نلاحظ أن القيمة المحسوبة أكبر تماما من القيمة المجدولة لتوزيع فيشر بنسبة معنوية $\alpha = 0.05$ و درجتي حرية $9 = \frac{31-9-4}{2} = \frac{n-m-2(k+1)}{2} = DF = \nu_1 = \nu_2$ أي $F_c > F_{0.05}(9,9)$. إذن نرفض H_0 و هذا يعني أن تباين الأخطاء غير متجانس.

للتأكد من النتيجة المتوصل إليها نستعمل إحصائية White لاختبار تجانس التباين. نقوم أولا بتقدير المعادلة الوسيطية، نحصل على:

$$\hat{\varepsilon}_t^2 = 13685.49 - 1.32X_t + 0.00013X_t^2$$
$$R^2 = 0.368; n = 31$$

حيث $\hat{\varepsilon}_t^2$ مربعات بواقي تقدير نموذج الادخار الكلي.

ثم نحسب إحصائية مضاعف لاغرانج، حيث:

$$LM = n \times R^2 = 31 \times 0.368 = 11.42$$

نلاحظ أن إحصائية مضاعف لاغرانج أكبر تماما من القيمة المجدولة لتوزيع χ^2 بنسبة معنوية $\alpha = 0.05$ و درجة حرية $DF = 2k = 2$ أي $LM > \chi^2_{0.05}(2) = 5.99$. إذن نرفض H_0 و هذا يعني أن تباين الأخطاء في هذه الحالة أيضا غير متجانس.

هناك اختبار آخر يهدف إلى معرفة مدى تجانس التباين الشرطي للأخطاء يسمى باختبار ARCH-LM. نقوم أيضا بتقدير المعادلة الوسيطية، نحصل على:

$$\hat{\varepsilon}_t^2 = 43402.09 + 0.27\hat{\varepsilon}_{i-1}^2$$
$$R^2 = 0.07711$$

حيث $\hat{\varepsilon}_t^2$ مربعات بواقي تقدير نموذج الادخار الكلي.

ثم نحسب إحصائية مضاعف لاغرانج، حيث:

$$LM = (n-1) \times R^2 = (31-1) \times 0.077 = 2.23$$

نلاحظ أن إحصائية مضاعف لاغرانج أصغر تماما من القيمة المجدولة لتوزيع χ^2 بنسبة معنوية $\alpha = 0.05$ و درجة حرية $DF = q = 1$ أي $LM < \chi^2_{0.05}(1) = 3.84$. إذن نقبل H_0 و هذا يعني أن التباين الشرطي للأخطاء متجانس. نستنتج أن التباين الهامشي للأخطاء غير متجانس أما التباين الشرطي فيعتبر ثابت.

الملحق:

أ. بالنسبة لمشكل الارتباط الذاتي، يمكن معالجته باستعمال RATS 5.04. نقدر النموذج بطريقة المربعات الصغرى:

```
linreg y / resids
#constant x
```

- اختبار Breusch-Godfrey:

```
linreg(noprint) resids
#constant x resids{1} resids{2}
compute BG_stat=%nobs*%rsquared
display BG_stat
```

- تقدير النموذج المصحح بطريقة دربين–واتسون المباشرة:

```
compute %durbin=0.446190
compute rho=1-%durbin/2
set dy / = y-rho*y{1}
set dx / = x-rho*x{1}
linreg dy / resids
#constant dx
```

- التقدير بطريقة Hidreth-Lu:

```
ar1(method=hilu) y
#constant x
```

- التقدير بطريقة Cochrane-Orcutt:

```
ar1(method=corc) y
#constant x
```

ب. بالنسبة لمشكل عدم تجانس التباين، يمكن معالجته أيضا باستعمال RATS 5.04.

نقدر نموذج الادخار بطريقة المربعات الصغرى:

```
linreg y / resids
#constant x
```

- اختبار White:

```
set sx / = x**2
set sresids / = resids**2
linreg(noprint) sresids
#constant x sx
compute white_stat=%nobs*%rsquared
display white_stat
```

- اختبار ARCH-LM:

```
linreg(noprint) sresids
#constant sresids{1}
compute archlm_stat=%nobs*%rsquared
display archlm_stat
```

تجدر الإشارة إلى أنه يمكن أيضا استعمال Eviews 4.0 فيعطي النتائج مباشرة وبطريقة بسيطة، فالمهتمون بعلم الاقتصاد القياسي و خاصة المبتدئون منهم يمكنهم استعمال هذا البرنامج باعتباره سهل الاستعمال. أما RATS 5.04 فيحتاج إلى معرفة اللغة المستعملة و كيفية كتابة التعليمات.

الفصل الرابع

طرق وتقنيات أخرى في تحليل الانحدار

الفصل الرابع

طرق وتقنيات أخرى في تحليل الانحدار

أوضحنا في الفصلين السابقين أهم التقنيات في النموذج الخطي العام، فلقد تطرقنا إلى صياغة مقدرات المربعات الصغرى و الكيفية التي تكون فيها هذه المقدرات خطية و غير متحيزة و تعديل هذه المقدرات في حالة الارتباط الذاتي بين الأخطاء و عدم تجانس التباين.

سندرس في هذا الفصل بعض الطرق الإضافية التي يمكن استخدامها في تحليل الانحدار المتعدد و بالتحديد معالجة حالة المتغيرات المبطأة بتعمق أكثر و تحليل النماذج غير الخطية حيث سنيتم تركيزنا على بعض الأدوات القياسية الأخرى للتقدير منها طريقة المربعات الصغرى غير الخطية و الطريقة غير المعلمية "طريقة النواة".

1. النماذج ذات المتغيرات المتباطئة زمنيا

عند بناء النماذج الاقتصادية من المهم أخذ عامل الزمن بعين الاعتبار، حيث نجد عادة وجود فترة زمنية بين حركة المتغيرات التابعة التي تستجيب للمتغيرات المستقلة. إن إدخال مثل هذه المتغيرات في تحليل الانحدار يجعل التحليل أشمل و أقرب إلى الواقع، حيث أن هناك متغيرات قد ترتبط بمتغيرات أخرى في نفس الفترة الزمنية كالنماذج الساكنة و في أغلب الحالات قد ترتبط بقيم ماضية لبعض المتغيرات فتصبح النماذج حركية.

ينبغي إدخال عامل التباطؤ الزمني للمتغير المستقل لأن في نماذج السلاسل الزمنية خاصة هناك فترة زمنية تقع بين اتخاذ القرار الاقتصادي و التأثير النهائي للتغير في متغير السياسة الاقتصادية و لا سيما إذا كان في فترة طويلة.

يحتل التباطؤ الزمني مكانا أساسيا في الاقتصاد، حيث يؤثر على طرق التحليل الاقتصادي سواء على المدى القصير أو الطويل. هناك أسباب تؤدي إلى وجود التباطؤ الزمني، منها أسباب نفسية بسبب العادات و التقاليد فقد لا يغير الناس عاداتهم

الاستهلاكية مباشرة بعد انخفاض الأسعار أو تزايد الدخل و هناك أيضا أسباب تقنية أو مؤسسية،.. الخ[1].

1.1. نماذج التخلف الزمني المتدرج

يطلق على هذا النوع من العلاقات فترات الإبطاء الموزعة، ويعني هذا أن المتغير التابع في أي فترة زمنية يعتمد على مجموع مرجح بالأوزان للمتغير المستقل في الفترات السابقة:

$$Y_t = b_0 + a_0 X_t + a_1 X_{t-1} + a_2 X_{t-2} + ... + a_k X_{t-k} + \varepsilon_t$$

حيث يتلاشى تأثير المتغير المستقل على المتغير التابع مع الوقت:

$$a_0 > a_1 > a_2 > ... > a_k$$

على سبيل المثال، الإنفاق الاستهلاكي قد يعتمد على مستويات الدخل المتاح الحالي و الدخل المتاح في فترات سابقة.

يمكن تبسيط النموذج باستعمال معامل التأخير L المعرف كما يلي: $LX_t = X_{t-1}$ و بصفة عامة: $L^i X_t = X_{t-i}$. لدينا:

$$Y_t = \sum_{i=1}^{k} a_i L^i X_t + b_0 + \varepsilon_t = \left[\sum_{i=1}^{k} a_i L^i \right] X_t + b_0 + \varepsilon_t$$

أي:

$$Y_t = A(L) X_t + b_0 + \varepsilon_t$$

حيث $A(L)$ كثير حدود من الدرجة k : $A(L) = a_0 + a_1 L + a_2 L^2 + ... + a_k L^k$

قد يكون عدد التباطؤ k محدودا أو غير محدود. بينما مجموع المعاملات a_i يؤول إلى نهاية محدودة. من جانب آخر، عندما تكون قيمة k غير معروفة فيمكن تحديد عدد التباطؤ وذلك بتصغير معياري Akaike و Schwarz المعرفين كما يلي:

$$AIC(k) = \ln\left(\frac{RSS_k}{n}\right) + \frac{2k}{n}$$

$$SC(k) = \ln\left(\frac{RSS_k}{n}\right) + \frac{k \ln n}{n}$$

1- وليد اسماعيل السيفو و أحمد محمد مشعل، 2003، ص 380.

حيث: RSS_k يمثل مجموع مربعات البواقي للنموذج المبطأ بالدرجة k

n يمثل حجم العينة

ln : اللوغاريتم النبيري.

وعلى الرغم من أن العلاقة تعد شكلا ملائما للتقدير في بعض الحالات، إلا أنها تسبب بعض المشاكل التي ترتبط بفترات الإبطاء الموزعة. نفقد مشاهدة لكل قيمة مبطأة إضافية للمتغير المستقل، إضافة إلى ذلك وجود عدد كبير من العالم التي ينبغي تقديرها و التي ستناظر متغيرات ترتبط بصورة قوية ببعضها البعض فسيجعل من الصعب عزل تأثير مختلف المتغيرات المستقلة على المتغير التابع. بمعنى آخر، ستكون تباينات المقدرات كبيرة. لمواجهة هذه المشاكل، اقترح بعض الإحصائيين نماذج إما أن تقلل من عدد المشاهدات التي تفقد بسبب الإبطاء أو تقلل من عدد المعالم التي ينبغي تقديرها.

النموذج الأول هو نموذج إبطاء كويك Koyck الذي يفترض تناقص الأوزان الخاصة بفترات الإبطاء هندسيا. ليكن λ عددا ثابتا تتراوح قيمته بين الصفر و الواحد حيث:

$$a_1 = \lambda a_0$$
$$a_i = \lambda^i a_0$$

لدينا: $$Y_t = b_0 + a_0 X_t + \lambda a_0 X_{t-1} + \lambda^2 a_0 X_{t-2} + ... + \lambda^k a_0 X_{t-k} + \varepsilon_t$$

أو: $$Y_t = b_0 + a_0 \left(X_t + \lambda X_{t-1} + \lambda^2 X_{t-2} + ... + \lambda^k X_{t-k} \right) + \varepsilon_t$$

الدالة $A(L)$ تكتب كما يلي:

$$A(L) = a_0 + a_0 \lambda L + a_0 \lambda^2 L^2 + ... + a_0 \lambda^k L^k$$

إذن النموذج $Y_t = A(L) X_t + b_0 + \varepsilon_t$ يمكن كتابته كما يلي:

$$B(L) Y_t = B(L) A(L) X_t + B(L) b_0 + B(L) \varepsilon_t$$

مع: $B(L) = A(L)^{-1}$

وعليه $A(L) = a_0 \left(1 + \lambda L + \lambda^2 L^2 + ...\right) = a_0 \left(1 - \lambda L\right)^{-1}$ يعتبر مجموع متتالية هندسية.

لدينا إذن: $$B(L) = \left(1 - \lambda L\right) / a_0$$

أي: $$\left(1 - \lambda L\right) Y_t = a_0 X_t + (1 - \lambda) b_0 + (1 - \lambda L) \varepsilon_t$$

أو:
$$Y_t = \lambda Y_{t-1} + a_0 X_t + (1-\lambda)b_0 + \varepsilon_t - \lambda\varepsilon_{t-1}$$

افترضنا أن الخطأ العشوائي $\nu_t = \varepsilon_t - \lambda\varepsilon_{t-1}$ يحقق الشروط كافة التي تفرضها الأخطاء العشوائية و لكن لسوء الحظ، فإن ذلك ليس صحيحا بالضرورة. فإذا كان ε_t في النموذج الأصلي يحقق الفرضيات الأساسية فإن ν_t في نموذج كويك Koyck لا يحققها عموما و بالتحديد الأخطاء العشوائية ν_t غير مستقلة ذاتيا، أي لا نتوقع أن تتحقق الفرضية $E(\nu_t\nu_{t-1}) = 0$ حيث:

$$E(\nu_t\nu_{t-1}) = E[(\varepsilon_t - \lambda\varepsilon_{t-1})(\varepsilon_{t-1} - \lambda\varepsilon_{t-2}] = E[\varepsilon_t\varepsilon_{t-1} - \lambda\varepsilon_t\varepsilon_{t-2} - \lambda\varepsilon_{t-1}^2 + \lambda^2\varepsilon_{t-1}\varepsilon_{t-2}] = -\lambda\sigma_\varepsilon^2 \neq 0$$

وهكذا فإن $\text{cov}(\nu_t, \nu_{t-1}) \neq 0$ و هذا يعني أن الأخطاء مرتبطة ذاتيا و هذا ما يتنافى مع إحدى فرضيات النموذج، كما أن $\text{cov}(\nu_t, Y_{t-1}) \neq 0$. نموذج كويك Koyck سيؤدي إلى انتهاك بعض الفرضيات، إضافة إلى ذلك مقدرات a_0، b_0 و λ تكون متحيزة و غير متسقة. يمكن التغلب على هذه الصعوبة عن طريق استخدام نموذج Solow وفق توزيع Pascal حيث معاملات نموذج الإبطاء توزع وفق:

$$a_i = (1-\lambda)^{r+1} C_{r+i}^i \lambda^i$$

حيث C_{r+i}^i يمثل معامل ذي حدين لـ Newton. ، r و λ هي المعالم مع $0 < \lambda < 1$ و $r \in N$

من أجل $r = 0$، النموذج هو التوزيع الهندسي لكويك Koyck.

النموذج العام يكتب كما يلي:

$$Y_t = b_0 + \sum_{i=0}^{\infty} (1-\lambda)^{r+1} C_{r+i}^i \lambda^i X_{t-i} + \varepsilon_t$$

أو:
$$Y_t = A(L)X_t + b_0 + \varepsilon_t$$

نفس البرهان السابق، نجد:

$$B(L)Y_t = B(L)A(L)X_t + B(L)b_0 + B(L)\varepsilon_t$$

مع: $B(L) = A(L)^{-1}$

من أجل $r = 0$، $B(L) = (1 - \lambda L)/a_0$

من أجل $r=1$ ، $B(L)=(1-\lambda L)^2/a_0$

لدينا: $$Y_t = 2\lambda Y_{t-1} - \lambda^2 Y_{t-2} + a_0 X_t + (1-2\lambda+\lambda^2)b_0 + v_t$$

حيث: $$v_t = (1-2\lambda L + \lambda^2 L^2)\varepsilon_t = \varepsilon_t - 2\lambda\varepsilon_{t-1} + \lambda^2\varepsilon_{t-2}$$

من أجل $r=2$ ، $B(L)=(1-\lambda L)^3/a_0$

لدينا:

$$Y_t = 3\lambda Y_{t-1} - 3\lambda^2 Y_{t-2} + \lambda^3 Y_{t-3} + a_0 X_t + (1-3\lambda+3\lambda^2-\lambda^3)b_0 + v_t$$

حيث:

$$v_t = (1-3\lambda L + 3\lambda^2 L^2 - \lambda^3 L^3)\varepsilon_t = \varepsilon_t - 3\lambda\varepsilon_{t-1} + 3\lambda^2\varepsilon_{t-2} - \lambda^3\varepsilon_{t-3}$$

لتقدير المعلم r ، اقترح Maddala and Rao (1971) استخدام طريقة المسح حيث دالة الهدف التي ينبغي تعظيمها هي معامل التحديد المصحح $\overline{R}^2$.

مثال 1:

لشرح هذا النوع من النماذج، نقوم بدراسة العلاقة بين معدل الفائدة طويل الأجل Y_i بدلالة معدل الفائدة قصير الأجل X_i في فترات سابقة. نقوم أولا بتحديد القيم المبطأة في النموذج وذلك بتصغير معياري Akaike و Schwarz ثم نقدر النموذج الأمثل. نأخذ معطيات المثال 2 في الفصل الثالث.

نقوم بحساب المعيارين بالاستعانة ببرنامج RATS 5.04:

```
CMOM
# CONSTANT SHORTRATE{0 TO 10} LONGRATE
DO MAXLAG=0,10
   LINREG(CMOM,NOPRINT) LONGRATE
   # CONSTANT SHORTRATE{0 TO MAXLAG}
   COMPUTE AKAIKE =%NOBS*LOG(%RSS)+%NREG*2.0
   COMPUTE SCHWARZ=%NOBS*LOG(%RSS)+%NREG*LOG(%NOBS)
   IF MAXLAG==0
      DISPLAY @4 'LAGS'  @20 'AKAIKE'  @35  'SCHWARZ'
   DISPLAY @5 #### MAXLAG  @20 ####.#### AKAIKE  @35 ####.#### SCHWARZ
END DO
```

النتائج مبينة في الجدول (2):

الجدول (1): نتائج البحث عن عدد القيم المبطأة المثلى

Schwarz	Akaike	القيم المبطأة
16.0111	13.2087	0
19.3527	15.1491	1
21.3068	15.7021	2
22.9842	15.9783	3
23.6774	15.2703	4
21.2458	11.4374	5
19.3088	8.0992	6
22.7055	10.0947	7
22.7523	8.7403	8
25.0222	9.6091	9
26.7826	9.9683	10

نلاحظ أن معيار AIC يأخذ القيمة الصغرى عند قيمة مبطأة k تساوي 6 و Schwarz أصغر ما يمكن عندما تكون القيمة المبطأة k مساوية إلى 0. إذا أخذنا بعين الاعتبار معيار AIC، فإن النموذج سيعاد تقديره بـ . 6 تباطؤات. يبين الجدول التالي نتائج التقدير كما يلي:

الجدول (2): نتائج التقدير

الاحتمال	قيم ستيودنت	الانحراف المعياري	المعامل المقدر	المتغير
0.0000	31.5048	0.2316	7.2989	الثابتة
0.8352	0.2105	0.0698	0.0147	X_t
0.4262	0.8106	0.0937	0.0760	X_{t-1}
0.6622	0.4427	0.0903	0.0399	X_{t-2}
0.9415	0.0742	0.0840	0.0062	X_{t-3}
0.8646	0.1725	0.0835	0.0144	X_{t-4}
0.5228	-0.6492	0.0824	-0.0535	X_{t-5}
0.0404	-2.0699	0.0653	-0.1353	X_{t-6}
مجموع مربعات البواقي: 0.7684			$R^2 = 0.46$	

نلاحظ أن لمعامل X_{t-6} معنوية إحصائية أي يختلف معنويا عن الصفر باعتبار أن إحصائية ستيودنت بالقيمة المطلقة أكبر من القيمة المجدولة عند مستوى معنوية 5% وهذا ما نلاحظه من خلال الاحتمال الذي يبقى أقل من 0.05 وهذا ما تؤكده النتائج السابقة الخاصة بالقيمة المبطأة المثلى.

بالاستعانة بمعطيات المثال (1)، نعيد تقدير النموذج السابق وفق نموذج الإبطاء لكويك و توزيع Pascal:

1- نفترض أن معاملات نموذج العلاقات المبطأة تتصاعد هندسيا، فتقدير معالم النموذج في هذه الحالة يتم باستعمال نموذج كويك. نتائج التقدير تظهر على الشكل التالي:

$$\hat{Y}_t = 1.9112 + 0.6954Y_{t-1} + 0.0349X_t$$
$$(2.6947) \quad (6.9139) \quad (2.7528)$$
$$n = 30$$
$$R^2 = 0.6552$$

حيث (.): قيم ستيودنت.

نحصل على:

$$\hat{\lambda} = 0.6954$$
$$\hat{a}_0 = 0.0349$$
$$\hat{b}_0 = 1.9112/(1-0.6954) = 6.2744$$

النموذج يكتب كما يلي:

$$\hat{Y}_t = 6.274 + 0.0349X_t + 0.6954 \times 0.0349X_{t-1} + (0.6954)^2 \times 0.0349X_{t-2} + \ldots$$

2 – نفترض الآن أن معاملات النموذج تتبع توزيع Pascal.

$$\hat{Y}_t = 2.10 + 0.93Y_{t-1} - 0.254Y_{t-2} + 0.026X_t$$
$$(2.988) \quad (5.124) \quad (-2.536) \quad (2.285)$$
$$n = 30$$
$$R^2 = 0.6839$$

حيث (.): قيم ستيودنت.

قمنا بتقدير نموذج الانحدار الذاتي من الدرجة الثانية و لكن نستطيع أن نقدر باستعمال درجات أكبر من 2. في مثالنا هذا اقتصرنا فقط على درجتين و السبب في ذلك يرجع إلى عدم معنوية معاملات Y_{t-3} و Y_{t-4}.

من خلال النتائج النظرية، لدينا:

$$\hat{\lambda} = \sqrt{0.254} \approx 0.5039$$
$$\hat{a}_0 = 0.026$$
$$\hat{b}_0 = 2.10/(1-2\hat{\lambda}+\hat{\lambda}^2) = 2.10/(1-1.0078+0.2539) = 8.5331$$

النموذج المقدر يكتب كما يلي:

$$\hat{Y}_t = 8.53 + \sum_{i=0}^{\infty}(1-0.50)^2 C_{1+i}^{i} 0.50^i X_{t-i}$$

2.1. نماذج الانحدار الذاتي الخطية

في هذا النوع من النماذج الزمنية، المتغير التابع Y_t يرتبط بـ k متغير مستقل $X_{i1},...,X_{ij},...,X_{ik}$ في القترة t و بقيم ماضية لنفس المتغير خلال الفترات السابقة $Y_{t-1},Y_{t-2},...,Y_{t-p}$، حيث النموذج يكتب كما يلي:

$$Y_i = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + ... + \beta_k X_{tk} + \varepsilon_t$$

أو أيضا:

$$Y_i = \sum_{i=1}^{p} \phi_j Y_{t-i} + X\beta + \varepsilon_t$$

حيث $X(n,k+1)$ هي مصفوفة المتغيرات المستقلة و $\beta(k+1,1)$ شعاع المعالم

في هذا النموذج، فرضية الاستقلالية بين المتغيرات المستقلة و الخطأ العشوائي غير محققة لأن المتغيرات $Y_{t-1},Y_{t-2},...,Y_{t-p}$ التي ترتبط بـ $\varepsilon_{t-1},\varepsilon_{t-2},...,\varepsilon_{t-p}$ عشوائية باعتبار أن Y_{t+1} دالة تابعة لـ Y_t الذي يرتبط بـ ε_t، حيث: $E(Y_{t+1}\varepsilon_t) \neq 0$. نلاحظ أنه إذا كانت المتغيرات المستقلة X_{tj} و الأخطاء ε_t مثبتة، فإن المتغيرات التابعة تعتبر حلولا لمعادلة التراجع:

$$Y_i = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + S_t$$

مع: $S_i = X\beta + \varepsilon_t$

أثبت الإحصائيون أن الحل العام لهذه المعادلة يعتبر غير مستقر إذا كان أحد جذور المعادلة المميزة $1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p=0$ خارج الدائرة الأحادية حيث B يسمى بمعامل التباطؤ أو التخلف الزمني. نعتبر إذن أن فرضية استقرارية السيرورة محققة.

نذكر فقط أن النموذج المذكور أعلاه لا يستعمل إلا نادرا، ففي أغلب الأحيان، نستعمل فقط نماذج الانحدار الذاتي من الدرجة الأولى و الذي يكتب رياضيا كما يلي:

$$Y_i = \phi Y_{t-1} + \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \ldots + \beta_k X_{tk} + \varepsilon_t$$

يكون هذا النموذج مستقرا إذا كان $|\phi| < 1$

إن طريقة التقدير المناسبة تعتمد على مدى ارتباط الأخطاء العشوائية ذاتيا. في حالة هذا النوع من النماذج المذكور أعلاه، لا يمكن الاستعانة باختبار دربين-واتسون لأن هذا الأخير غير فعال و متحيز و لهذا اقترح Durbin (1970) إحصائية أخرى تسمى بإحصائية Durbin « h » و التي تعرف رياضيا كما يلي:

$$h = \hat{\rho}\sqrt{\frac{n}{1-n\hat{\sigma}^2_{\hat{\phi}}}}$$

حيث $\hat{\rho} = 1 - DW/2$ (DW تعبر عن إحصائية دربين-واتسون)، n حجم العينة و $\hat{\sigma}^2_{\hat{\phi}}$ التباين المقدر للمعامل ϕ.

تتوزع إحصائية Durbin « h » بصفة مقاربة وفق التوزيع الطبيعي المعياري:

$$H_0 : \rho = 0 \quad H_0 : h = 0$$
$$H_1 : \rho \neq 0 \quad H_1 : h \neq 0$$

إذا كان $|h| \leq t_{\alpha/2}$ فإننا نقبل الفرضية H_0 أي هناك استقلالية تامة بين الأخطاء ($t_{\alpha/2}$ هي القيمة المجدولة للتوزيع الطبيعي عند مستوى معنوية α)

نلاحظ أنه إذا كان $n\hat{\sigma}^2_{\hat{\phi}} \geq 1$ فمن المستحيل استعمال هذه الإحصائية و عليه نستعين بإحصائية دربين-واتسون الكلاسيكية لاختبار استقلالية الأخطاء مع الأخذ بعين الاعتبار مناطق الشك.

في حالة غياب الارتباط الذاتي بين الأخطاء، نستعمل طريقة المربعات الصغرى العادية لتقدير معالم نموذج الانحدار الذاتي من الدرجة الأولى و المقدرات المتحصل عليها غير متحيزة بصفة مقاربة و ذات أصغر تباين. أما في العينات الصغيرة، عند تقدير نموذج الانحدار الذاتي من الدرجة p ، نتائج التقدير لن تكون جيدة لأن عدد المشاهدات و التي من خلالها تمت عملية التقدير هي $n-p$. إضافة إلى ذلك، لا تسمح مشاكل الارتباط الخطي بين المتغيرات المستقلة التي يمكن مصادفتها بتطبيق طريقة المربعات الصغرى العادية و عليه لا ينبغي استخدام هذه الطريقة إلا إذا كان حجم العينة كبيرا و تطبيقيا يجب أن يتعدى حجم العينة 15 مشاهدة.

مثال 2:

للتوضيح أكثر، نقوم باختبار العلاقة بين الأسعار الرسمية لطن من القهوة Y_t و الأسعار المطبقة على الصادرات من قبل الدول المنتجة للقهوة X_t خلال الفترة الممتدة بين سنتي 1993 و 2008. نقترح تقدير النموذج التالي:

$$Y_t = \beta_0 + \beta_1 X_t + \phi Y_{t-1} + \varepsilon_t$$

و هذا يعني أن السعر الرسمي لطن من القهوة يرتبط بالسعر الرسمي للسنة الماضية و بالسعر المطبق على الصادرات من قبل الدول المنتجة. لدينا المعطيات التالية الموضحة في الجدول (4):

الجدول (3): السعر الرسمي و السعر المطبق على الصادرات لطن من القهوة

X_t	Y_t	السنة
615	455	1993
665	500	1994
725	555	1995
795	611	1996
870	672	1997
970	748	1998
1095	846	1999
1235	954	2000
1415	1090	2001

X_t	Y_t	السنة
1615	1243.5	2002
1795	1390	2003
2015	1559	2004
2315	1781	2005
2660	2046.5	2006
2990	2311	2007
3280	2551	2008

نقدر النموذج المقترح بطريقة المربعات الصغرى، نحصل على النتائج التالية باستعمال برنامج الكمبيوتر RATS 5.04:

```
Linear Regression - Estimation by Least Squares
Dependent Variable Y
Annual Data From 1994:01 To 2008:01
Usable Observations     15     Degrees of Freedom    12
Centered R**2     0.999980      R Bar **2   0.999977
Uncentered R**2   0.999996      T x R**2      15.000
Mean of Dependent Variable      1257.2000000
Std Error of Dependent Variable  663.6160088
Standard Error of Estimate         3.1985455
Sum of Squared Residuals         122.76832042
Regression F(2,12)               301313.1129
Significance Level of F           0.00000000
Durbin-Watson Statistic             0.673643

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant               -7.204279236  1.829761263     -3.93728  0.00197257
2.  Y{1}                    0.222825530  0.036615691      6.08552  0.00005455
3.  X                       0.623202236  0.025194776     24.73537  0.00000000
```

قدرت هذه المعادلة على 15 مشاهدة $(n-1)$ لأن النموذج يحتوي على متغير مبطأ Y_{t-1}. نلاحظ أن للنموذج قدرة تفسيرية عالية جدا $R^2 = 0.9999$ و لجميع المعالم معنوية إحصائية أي تختلف كلها معنويا عن الصفر بنسبة معنوية 5% لأن قيم ستيودنت بالقيمة المطلقة أكبر تماما من القيمة الحرجة لتوزيع ستيودنت بدرجة حرية 12 و نسبة معنوية

5%. إضافة إلى ذلك، معامل Y_{t-1} المقدر بالقيمة المطلقة أقل تماما من الواحد و هذا يعني أن النموذج مستقر.

إلا أن إحصائية دربين–واتسون تشير إلى وجود ارتباط ذاتي بين الأخطاء. للتأكد من ذلك، علينا إذن استعمال إحصائية Durbin « h » لهذا الغرض:

$$h = \hat{\rho}\sqrt{\frac{n}{1-n\hat{\sigma}^2_{\hat{\phi}}}} = 0.66318\sqrt{\frac{15}{1-15\times 0.00134}} = 2.56676 > t_{0.05} = 1.96$$

نرفض الفرضية H_0 أي هناك ارتباط ذاتي بين الأخطاء.

نقوم الآن بتقدير النموذج على المتغيرات ذات الفروقات من الدرجة الأولى و الهدف من ذلك هو المقارنة بين النموذج الأصلي و النموذج الجديد (الفروقات من الدرجة الأولى) و في هذه الحالة سنفقد مرة أخرى مشاهدة أخرى عند حساب الفروقات من الدرجة الأولى. لتكن النتائج التالية:

```
Linear Regression - Estimation by Least Squares
Dependent Variable DY
Annual Data From 1995:01 To 2008:01
Usable Observations     14     Degrees of Freedom    11
Centered R**2     0.999234      R Bar **2   0.999095
Uncentered R**2   0.999846      T x R**2      13.998
Mean of Dependent Variable      146.50000000
Std Error of Dependent Variable  76.39497567
Standard Error of Estimate        2.29804607
Sum of Squared Residuals         58.091173217
Regression F(2,11)                 7177.8245
Significance Level of F           0.00000000
Durbin-Watson Statistic             1.969185

   Variable                     Coeff        Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                2.8441890605 1.3577456191      2.09479  0.06014044
2.  DY{1}                   0.2387855569 0.0251813152      9.48265  0.00000125
3.  DX                      0.5996158163 0.0193807615     30.93871  0.00000000
```

قدرت هذه المعادلة التي تتضمن المتغيرات المحولة عن طريق الفروقات من الدرجة الأولى على 14 مشاهدة $(n-2)$. نلاحظ أن لهذا النموذج قدرة تفسيرية عالية جدا

$R^2 = 0.9992$ و لجميع المعالم معنوية إحصائية أي تختلف كلها معنويا عن الصفر بنسبة معنوية 5% لأن قيم ستيودنت بالقيمة المطلقة تبقى دائما أكبر تماما من القيمة الحرجة لتوزيع ستيودنت بدرجة حرية 11 و نسبة معنوية 5% إلا أن إحصائية دربين-واتسون تشير هذه المرة إلى استقلالية تامة بين الأخطاء.

نلاحظ أن الفرق بين معاملات النموذج الثاني (على المتغيرات ذات الفروقات من الدرجة الأولى) و الأول ضئيل: (0.22 و 0.23) بالنسبة للمعامل الأول و (0.62 و 0.59) بالنسبة للمعامل الثاني.يمكن اعتبار النتائج المتحصل عليها في النموذج الأول مقبولة. ومع ذلك نقوم بتصحيح النموذج من الارتباط الذاتي بين الأخطاء وذلك باستعمال طريقة Cochrane-Orcutt للتقدير. النتائج تظهر على مخرجات برنامج RATS 5.04:

```
Regression with AR1 - Estimation by Cochrane-Orcutt
Dependent Variable Y
Annual Data From 1995:01 To 2008:01
Usable Observations      14     Degrees of Freedom    10
Centered R**2      0.999997      R Bar **2   0.999996
Uncentered R**2    0.999999      T x R**2      14.000
Mean of Dependent Variable      1311.2857143
Std Error of Dependent Variable  653.4583359
Standard Error of Estimate        1.3070776
Sum of Squared Residuals        17.084519106
Durbin-Watson Statistic            2.278697
Q(3-1)                             4.371054
Significance Level of Q           0.11241850

   Variable                     Coeff        Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                -3.170642272  1.343751037     -2.35955  0.03998422
2.  Y{1}                     0.232100312  0.015792767     14.69662  0.00000004
3.  X                        0.615141751  0.010986171     55.99237  0.00000000
*******************************************************************************
4.  RHO                      0.391048353  0.121682865      3.21367  0.00927417
```

يظهر جليا أن لكل المعالم معنوية إحصائية و للنموذج قدرة تفسيرية ممتازة. يمكن القول إذن أن الأسعار الرسمية في الفترة الحالية تتأثر أكثر بالأسعار الحالية المطبقة على

الصادرات من قبل الدول المنتجة و ليس بالأسعار الرسمية في الفترة السابقة و هذا ما نلاحظه من خلال قيم ستيودنت.

3.1. أمثلة عن النماذج الحركية

هناك نماذج أخرى يمكن أن توصلنا إلى معادلة كويك و منها نموذج التصحيح الجزئي و نموذج التوقعات المكيفة. وعليه لا بد من إعطاء فكرة عن هذا النوع من النماذج.

1.3.1. نموذج التصحيح الجزئي The Partial Adjustment Model

و يسمى في بعض الأحيان بنموذج تعديل المخزون The Stock Adjustment Model. نفترض مثلا حالة التوازن على المدى الطويل مع وجود كمية من رأس المال المخزون تستعمل للحصول على كمية من الإنتاج في ظل التقدم العلمي السائد و سعر الفائدة و للتبسيط أكثر نفترض المستوى من رأسمال يساوي Y_t^D و هو دالة خطية لمستوى الإنتاج X_t وفق النموذج التالي:

$$Y_t^D = \alpha + \beta X_t + \varepsilon_t$$

وأن العلاقة بين المستوى الفعلي للمتغير التابع و المستوى المرغوب تعبر عنها رياضيا كما يلي:

$$Y_t - Y_{t-1} = \lambda(Y_t^D - Y_{t-1})$$

حيث λ تعبر عن معامل التعديل Coefficient of Adjustment مع $0 \leq \lambda \leq 1$ أما $Y_t - Y_{t-1}$ هو التغير الفعلي و $Y_t^D - Y_{t-1}$ التغير المرغوب.

تتضمن هذه المعادلة الحركة الجزئية من موقع الأساس λ إلى الموقع الأمثل. كلما كان λ يقترب من الواحد، كلما كبر التعديل في الفترة الجارية. بدمج المعادلتين الأخيرتين، نحصل على:

$$Y_t - Y_{t-1} = \lambda(\alpha + \beta X_t + \varepsilon_t - Y_{t-1})$$

أي:

$$Y_t - Y_{t-1} = \lambda\alpha + \lambda\beta X_t + \lambda\varepsilon_t - \lambda Y_{t-1}$$

وعليه:

$$Y_t - Y_{t-1} = \lambda\alpha + \lambda\beta X_t + \lambda\varepsilon_t - \lambda Y_{t-1}$$

و في الأخير نحصل على: $Y_t = (1-\lambda)Y_{t-1} + \lambda\alpha + \lambda\beta X_t + \lambda\varepsilon_t$

نضع $\beta_0 = \lambda\alpha$ ، $\beta_1 = \lambda\beta$ ، $\phi = 1-\lambda$ و $u_t = \lambda\varepsilon_t$. إذن لدينا:

$$Y_t = \phi Y_{t-1} + \beta_0 + \beta_1 X_t + u_t$$

يمكن استعمال طريقة المربعات الصغرى العادية لتقدير معالم هذا النموذج.

2.3.1. نموذج التوقعات المكيفة The Adaptive Expectation Model

في هذا النموذج، قيم المتغير التابع Y_t دالة تابعة لقيم المتغير المستقل المتوقعة، حيث:

$$Y_t = \alpha + \beta X_t^p + \varepsilon_t$$

حيث X_t^p هي القيمة المتوقعة للمتغير المستقل.

على سبيل المثال، الإنتاج في مؤسسة ما دالة تابعة للقيم المتوقعة للمبيعات أو كمية النقود دالة تابعة لسعر الفائدة المتوقع. لا يمكن استخدام هذه المعادلة مباشرة للتقدير لأن X_t^p غير معلومة أي لا تتوفر عنها بيانات و عليه فلا بد من تحقق بعض الفرضيات المتعلقة بصياغة التوقعات و الفرضية العامة هي التوقعات المكيفة و التي تعرف رياضيا وفق العلاقة التالية:

$$X_t^p - X_{t-1}^p = \lambda(X_t - X_{t-1}^p)$$

حيث تشير λ إلى معامل التوقع Coefficient of Expectation مع $0 \le \lambda \le 1$. تعرف هذه الفرضية بفرضية التوقع المتطور Progressive Expectation أو فرضية تعلم الخطأ[1] Error Learning Hypothesis.

تبين المعادلة الأخيرة أن التوقعات تتكون من الأجزاء التي تضيفها λ كل فترة زمنية إلى أن تسد الثغرة بين القيمة الحالية للمتغير و قيمته المتوقعة سابقا. هذه المعادلة يمكن كتابتها كما يلي:

$$X_t^p = \lambda X_t + (1-\lambda)X_{t-1}^p$$

و بنشر ها نحصل على: $X_t^p = \lambda X_t + (1-\lambda)X_{t-1} + \lambda(1-\lambda)^2 X_{t-2} + \dots$

1- اقترح هذه الفرضية كل من (1956) Gagan و (1957) Friedman.

أي: $X_t^p = \lambda \sum_{i=0}^{\infty} (1-\lambda)^i X_{t-i}$

نعوض X_t^p في المعادلة $Y_t = \alpha + \beta X_t^p + \varepsilon_t$، نحصل على العلاقة التالية:

$$Y_t = \alpha + \beta\lambda \sum_{i=0}^{\infty} (1-\lambda)^i X_{t-i} + \varepsilon_t$$

و التي تعبر عن نموذج التخلف الزمني المتدرج حيث تحويل كويك Koyck يسمح بكتابته على شكل انحدار ذاتي:

$$Y_t = \lambda\alpha + \lambda\beta X_t + (1-\lambda)Y_{t-1} + \left[\varepsilon_t - (1-\lambda)\varepsilon_{t-1}\right]$$

يمكن تقدير نموذج الانحدار الذاتي في حالة الارتباط الذاتي بين الأخطاء من الدرجة الأولى حيث نحصل على مقدرات لكل من λ، α و β.

مثال 3:

نأخذ معطيات المثال 1 على معدلات الفائدة قصيرة و طويلة الأجل. نقدر نموذج الانحدار الذاتي من الدرجة الأولى مع وجود ارتباط ذاتي بين الأخطاء (طريقة Hidreth-Lu):

$$\hat{Y}_t = \underset{(3.92)}{4.54} + \underset{(2.63)}{0.27}Y_{t-1} + \underset{(2.50)}{0.09}X_t$$

$$R^2 = 0.65, \quad n = 28, \quad DW = 1.79, \quad \hat{\rho} = 0.58$$

لدينا:

$$\hat{\lambda} = 1 - 0.27 = 0.73$$
$$\hat{\beta} = 0.09 / 0.73 = 0.12$$
$$\hat{\alpha} = 4.54 / 0.73 = 6.21$$

و عليه نموذج التصحيح (التعديل) الجزئي المقدر هو:

$$\hat{Y}_t^D = 6.21 + 0.12X_t$$

مع:

$$Y_t - Y_{t-1} = 0.73(Y_t^D - Y_{t-1})$$

أما نموذج التوقعات المكيفة المقدر هو:

$$\hat{Y}_t = 6.21 + 0.12X_t^p$$

مع: $X_t^p - X_{t-1}^p = 0.73(X_t - X_{t-1}^p)$

2. النماذج غير الخطية

لقد افترضنا أن شكل العلاقة التي رغبنا في تقديرها هو الشكل الخطي، ففي الحقيقة الشكل الخطي يعد شرطا مقيدا جدا وعادة ما تقترح النظرية الاقتصادية أو شكل انتشار النقط المشاهدة أن العلاقة بين المتغيرات غير خطية. و التساؤل هنا كيف يمكن التعامل مع العلاقات غير الخطية؟ هناك نوعان من النماذج: نماذج غير خطية يمكن تحويلها إلى شكل خطي و نماذج غير خطية لا يمكن تحويلها إلى شكل خطي.

1.2. التحويل الخطي للنماذج غير الخطية:

قبل تقدير العلاقة بين المتغير التابع و المتغيرات المستقلة، يجب أولا البحث عن أنسب الصيغ الرياضية التي تعبر عن هذه العلاقة تعبيرا دقيقا ولتحقيق ذلك يجب التعرف على الشكل البياني الحقيقي للعلاقة بين المتغيرات، ويتم ذلك بواسطة النظرية الاقتصادية أو الدراسات التطبيقية السابقة أو الرسم البياني للمتغير التابع وكل متغير مستقل على حدا ثم اختيار أنسب الصيغ الرياضية التي تتلاءم مع الشكل البياني الحقيقي للعلاقة محل الدراسة.

مثال 4:

لنأخذ مثالا عن دالة كوب-دوغلاس التي تعطي العلاقة بين الإنتاج و عوامله، نأخذ عينة من 25 مؤسسة لتقدير الإنتاج بدلالة العمل L و رأس المال K المعطاة بالعلاقة التالية:

$$Q_i = \beta_0 K_i^{\beta_1} L_i^{\beta_2} \varepsilon_i \,, \quad i = 1,\ 2,\ \ldots,\ 25$$

حيث ε_i: الخطأ العشوائي

تعتبر هذه الدالة غير خطية من حيث المعالم لكن يمكن تحويلها إلى شكل خطي و ذلك عن طريق التحويل اللوغاريتمي. النموذج المراد تقديره هو كالتالي:

$$\ln Q_i = \ln \beta_0 + \beta_1 \ln K_i + \beta_2 \ln L_i + \ln \varepsilon_i$$

$$Q_i^* = \beta_0^* + \beta_1 K_i^* + \beta_2 L_i^* + \varepsilon_i$$

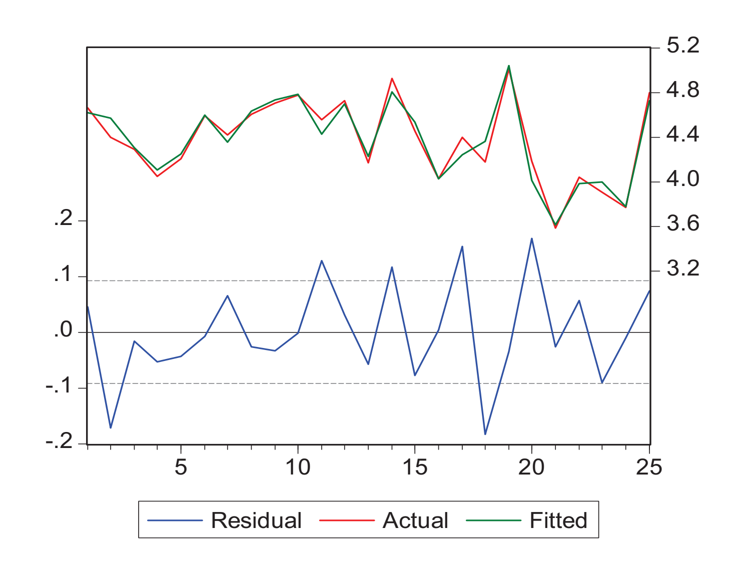
حيث L_i^* و K_i^* هما المتغيران المستقلان و Q_i^* المتغير التابع أما ε_i فهو الخطأ العشوائي. نستعمل طريقة المربعات الصغرى لهذا الغرض، فنحصل على النتائج التالية باستعمال Eviews 5.0:

Dependent Variable: LOGQ
Method: Least Squares
Date: 05/22/09 Time: 03:09
Sample: 1 25
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.480712	0.128732	19.27042	0.0000
LOGL	0.257384	0.026983	9.538866	0.0000
LOGK	0.640170	0.034761	18.41619	0.0000

R-squared	0.941473	Mean dependent var	4.375637
Adjusted R-squared	0.936152	S.D. dependent var	0.365297
S.E. of regression	0.092304	Akaike info criterion	-1.815293
Sum squared resid	0.187440	Schwarz criterion	-1.669028
Log likelihood	25.69117	F-statistic	176.9458
Durbin-Watson stat	2.545210	Prob(F-statistic)	0.000000

الشكل رقم (1): التمثيل البياني للقيم الحقيقية و المقدرة مع بواقي التقدير



نلاحظ أن للنموذج قدرة تفسيرية عالية جدا لأن لوغاريتم العمل و لوغاريتم رأس المال يفسران لوغاريتم الإنتاج بنسبة 93.61% و هذا ما نلاحظه من خلال الشكل (1) حيث أن الإنتاج المقدر يقترب من القيم المشاهدة. يمكن التأكد من ذلك باستعمال إحصائية فيشر التي تساوي إلى 176.94 فهي أكبر تماما من القيمة الحرجة لتوزيع فيشر بدرجتي حرية 2 و 22 و هذا يعني أننا نرفض الفرضية H_0. إضافة إلى ذلك، لمعالم النموذج معنوية إحصائية أي تختلف كلها معنويا عن الصفر بنسبة دلالة 5% باعتبار أن إحصائيات ستيودنت أكبر تماما من القيمة الحرجة لنفس التوزيع بدرجة حرية 22: $t_{0.05} = 2.070$. الإنتاج المقدر يكتب كما يلي:

$$\hat{Q}_i = 11.94 K_i^{0.6401} L_i^{0.2573}$$

حيث يتم حساب الثابتة المقدرة كما يلي: $\hat{\beta}_0 = e^{2.4807} = 11.94$

هناك أشكال أخرى من النماذج منها دالة القطع المكافئ حيث أن النموذج يأخذ الشكل $Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i,\ i = 1,.....,n$ فهي دالة خطية لـ Y_i بدلالة X_i و X_i^2 أي أنها غير خطية بالنسبة للمتغيرات و خطية بالنسبة للمعالم، فعلى سبيل المثال التكاليف الكلية دالة تابعة للإنتاج حيث:

$$TC_i = \beta_0 + \beta_1 Q_i + \beta_2 Q_i^2 + \varepsilon_i,\ i = 1,.....,n$$

حيث TC_i التكاليف الكلية و Q_i الإنتاج أما ε_i فهو الخطأ العشوائي. فيكفي إذن تقدير هذا النوع من النماذج بطريقة المربعات الصغرى العادية.

2.2. طرق تقدير النماذج غير الخطية التي غير قابلة للتحويل إلى شكل خطي

هناك بعض النماذج غير الخطية، من المستحيل تحت فرضيات معينة تطبيق طريقة المربعات الصغرى. نستعمل تقنيات أخرى للتقدير مهما يكن نوع الخوارزمية المستعملة[1].

يمكن تحويل هذه النماذج إلى الشكل الخطي عن طريق نشر تايلور Taylor وذلك بإعطاء قيم ابتدائية للمعالم ويتم تقديرها عن طريق تكرار العملية iteration. تستخدم

1- لمزيد من التفاصيل أنظر Greene, W.H. (2000), chapitre 10

طريقة المربعات الصغرى على المعادلة التي تم تحويلها إلى شكل خطي من أجل تقدير معاملات جديدة. تسمح هذه الأخيرة عن طريق نشر جديد محدود بتحويل خطي جديد، يتم إيقاف العملية عندما تكون المعاملات ساكنة نسبيا من مرحلة إلى مرحلة أخرى.

لكي تكون هذه الطريقة فعالة، ينبغي أن تكون القيم الابتدائية للمعالم قريبة من القيم المثلى. إذا لم تكن كذلك فإن التقدير غير جيد أي لا يوجد تقارب Convergence.

ليكن النموذج غير الخطي التالي: $Y_i = f(X,\beta) + \varepsilon_i$ حيث X هي مصفوفة المتغيرات المفسرة بعدها $(n, k+1)$ و β شعاع المعالم ذو بعد $k+1$.

في ظل الفرضيات الأساسية الكلاسيكية للنموذج، يمكن إيجاد مقدر لشعاع المعالم و ذلك بتصغير مجموع مربعات البواقي:

$$S(\beta) = \hat{\varepsilon}'\hat{\varepsilon} = \left[y_t - f(X,\beta)\right]'\left[y_t - f(X,\beta)\right]$$

لدينا $k+1$ مشتق جزئي من الدرجة الأولى:

$$\frac{\partial S}{\partial \beta} = -2\frac{\partial f(X,\beta)}{\partial \beta}\left[y_t - f(X,\beta)\right] = 0$$

حيث:

$$\frac{\partial f(X,\beta)}{\partial \beta} = Z(\beta) = \begin{pmatrix} \dfrac{\partial f(x_1,\beta)}{\partial \beta_0} & \cdots & \dfrac{\partial f(x_1,\beta)}{\partial \beta_k} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f(x_n,\beta)}{\partial \beta_0} & \cdots & \dfrac{\partial f(x_1,\beta)}{\partial \beta_k} \end{pmatrix}$$

لتكن $Z(\beta^{(1)})$ هذه المصفوفة المحسوبة من أجل القيم الخاصة لـ β . $(\beta = \beta^{(1)})$.ثم باستعمال نشر محدود لتايلور بجوار $\beta^{(1)}$، يمكن تقريب المشاهدة:

$$f(x_t,\beta) \approx f(x_t,\beta^{(1)}) + \left[\left.\frac{\partial f(x_t,\beta)}{\partial \beta_0}\right|_{\beta=\beta^{(1)}} \quad \cdots \quad \left.\frac{\partial f(x_t,\beta)}{\partial \beta_k}\right|_{\beta=\beta^{(1)}} \right]\left(\beta - \beta^{(1)}\right)$$

بعبارة أخرى:

$$f(X,\beta) \approx f(X,\beta^{(1)}) + Z(\beta^{(1)})\left(\beta - \beta^{(1)}\right)$$

إذن: $Y \approx f(X,\beta^{(1)}) + Z(\beta^{(1)})\left(\beta - \beta^{(1)}\right) + \varepsilon$

أي:

$$Y \approx f(X,\beta^{(1)}) + Z(\beta^{(1)})\beta - Z(\beta^{(1)})\beta^{(1)} + \varepsilon$$

نضع: $Y^*(\beta^{(1)}) = Y - f(X,\beta^{(1)}) + Z(\beta^{(1)})\beta^{(1)}$

نستطيع من خلال العلاقة الأخيرة أن نقرب النموذج غير الخطي إلى شكل خطي[1]:

$$\widetilde{Y}(\beta^{(1)}) = Z(\beta^{(1)})\beta + \varepsilon$$

يمكن تقدير معلم هذا النموذج الخطي كما يلي:

$$\begin{aligned}\beta^{(2)} &= \left[Z(\beta^{(1)})' Z(\beta^{(1)})\right]^{-1} Z(\beta^{(1)})' \widetilde{Y}(\beta^{(1)}) \\ &= \beta^{(1)} + \left[Z(\beta^{(1)})' Z(\beta^{(1)})\right]^{-1} Z(\beta^{(1)})' \left[Y - f(X,\beta^{(1)})\right]\end{aligned}$$

والذي يعطي أيضا قيم جديدة للشعاع $\beta = \beta^{(2)}$ ويمكن مواصلة عملية البحث عن القيم المثلى لشعاع المعالم إلى غاية p إعادة، فنلاحظ سكونا نسبيا للمعاملات المقدرة: $\hat{\beta} = \beta^{(p)} = \beta^{(p-1)}$.

تجدر الإشارة إلى أن طريقة التقدير المتبعة لا تكون فعالة إلا إذا تمكنا من اختيار القيم الابتدائية للمعاملات بشكل ممتاز وفق النمذجة الاقتصادية و المعطيات.

مثال 5:

المطلوب دراسة العلاقة بين الاستهلاك الإجمالي C_t و الدخل المتاح Y_{dt} خلال الفترة الممتدة من الفصل الأول لعام 1947 إلى غاية الفصل الثالث من سنة 1995 و فق النموذج التالي:

$$C_t = a_0 + a_1 Y_{dt}^{a_2} + \varepsilon_t$$

نلاحظ أن هذه العلاقة لا يمكن تحويلها إلى شكل خطي و عليه يتم تقدير معالم هذا النموذج بطريقة المربعات الصغرى غير الخطية اعتمادا على تقنية Gauss-Newton. نتائج التقدير تظهر في مخرجات برمجية RATS 5.04:

1- تم تقريب النموذج غير الخطي إلى دالة خطية باستعمال خوارزمية Gauss-Newton والتي تعتمد على النشر المحدود لتايلور.

```
Nonlinear Least Squares - Estimation by Gauss-Newton
Convergence in    41 Iterations. Final criterion was  0.0000042 <  0.0000100
Dependent Variable C
Quarterly Data From 1947:01 To 1995:03
Usable Observations    195     Degrees of Freedom   192
Centered R**2     0.999019      R Bar **2   0.999009
Uncentered R**2   0.999841      T x R**2     194.969
Mean of Dependent Variable      1974.4145752
Std Error of Dependent Variable  870.2600021
Standard Error of Estimate        27.3989774
Sum of Squared Residuals        144135.16055
Regression F(2,192)               97763.0676
Significance Level of F           0.00000000
Durbin-Watson Statistic             0.597398

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  A0                       256.33157074  16.71186748     15.33830  0.00000000
2.  A1                         0.19519987   0.02109505      9.25335  0.00000000
3.  A2                         1.17995069   0.01263515     93.38638  0.00000000
```

النموذج المقدر يكتب كما يلي: $\hat{C}_t = 256.33 + 0.19Y_{dt}^{1.17}$ إذن يمكن قبول هذا النموذج إحصائيا معتمدا على النتائج المتحصل عليها (للمعالم معنوية إحصائية، للنموذج قدرة تفسيرية عالية جدا) إلا أن هناك ارتباط ذاتي بين الأخطاء و هذا ما نلاحظه من خلال إحصائية دربين-واتسون و عليه لا بد من تصحيح النموذج من الارتباط الذاتي.

نذكر فقط أنه لتقدير هذا النوع من النماذج استعنا بالبرنامج التالي:

```
NONLIN A0 A1 A2
FRML CONSFRML C = A0 + A1*YD**A2
COMPUTE A0=A1=A2=1.0
NLLS(FRML=CONSFRML,ITERS=50) C
```

3. مدخل إلى الانحدار اللا معلمي

يعتمد هذا النوع من النماذج على تقنيات حرة "لا معلمية" فهي لا تفترض وجود عائلة نماذج بل ترتكز على الفكرة "أترك المعطيات تتكلم وحدها" " Let the data speak for themselves". عرف الإحصاء غير المعلمي تطورا كبيرا في الآونة الأخيرة و لا سيما في دراسة العلاقات الاقتصادية حيث لم يستخدم في السلاسل الزمنية إلا في بداية

الثمانينات[1]. مبدأ هذه الطريقة يتمثل في دراسة معالم تمركز أو تشتت التوزيع الشرطي للظاهرة Y علما بالمتغير $X = x$.

يتم اختيار هذا الطريقة عند تعميم الانحدار غير الخطي المعلمي، فعندما تكون هناك صعوبة في اختيار الشكل الملائم للنموذج غير الخطي، نستعين بتقنيات حرة لتقدير الظاهرة الاقتصادية معتمدا على المتغيرات المستقلة. إضافة إلى ذلك، التقدير غير المعلمي للانحدار لا يعتمد على أي معلم أي لا يتعلق الأمر بتقدير شعاع معالم ذي بعد محدود بل عدد المعالم المراد تقديرها غير منتهي.

1.3. تقدير التوقع الرياضي الشرطي بطريقة النواة

لنعتبر النموذج غير الخطي التالي:

$$Y_i = f(X_i) + \varepsilon_i$$

حيث Y_i هو المتغير المُفَسَّر أو التابع و X_i المتغير المُفَسِّر أو المستقل و ε_i الخطأ العشوائي الذي يحقق الفرضيات الكلاسيكية للنموذج.

يمثل المتغير ε_i الخطأ في تفسير Y_i، ومنه يمكن كتابته انطلاقا من العلاقة $\varepsilon_i = Y_i - f(X_i)$ ويرجع سبب وجود الخطأ العشوائي إلى إهمال بعض المتغيرات المستقلة التي يمكن أن تؤثر على المتغير التابع في النموذج أو الصياغة الرياضية غير السليمة أو حدوث خطأ في كل من تجميع البيانات وقياس المتغيرات الاقتصادية. إذن يكون المتغير التابع دالة غير خطية (لا معلمية) للمتغير المستقل مضافا إليه حد الخطأ.

يتم تقدير هذا النموذج بطريقة تسمى بطريقة النواة Kernel method. أدخلت هذه الطريقة لأول مرة من طرف Rosenblatt (1956) لتقدير دالة الكثافة ثم أخذت ثانية من طرف Nadaraya (1964) وWatson (1964) لتقدير دالة الانحدار. الهدف إذن هو تقدير $E(Y \mid X = x)$ انطلاقا من الزوج (X_i, Y_i).

1- أنظر Bosq et Lecoutre (1982)، Härdle (1990).

ليكن μ القانون الذي يخضع له الزوج (X,Y)، $\hat{\mu}$ مقدره المعرف بالصيغة:

$$\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}\delta_{(X_i,Y_i)}$$

و υ قانون X و $\hat{\upsilon}$ مقدره المعرف كما يلي:

$$\hat{\upsilon}=\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}$$

نفترض أن للمتغير X دالة كثافة $g\neq 0$. يمكن كتابة $f(x)$ كما يلي:

$$f(x)=\frac{\int yd\mu(x,y)}{g(x)}$$

بوضع $\varphi(x)=\int yd\mu(x,y)$، الدالة $f(x)$ تكتب على الشكل:

$$f(x)=\frac{\varphi(x)}{g(x)}$$

تعرف المقدرات غير المعلمية $\hat{g}$ لـ g. و $\hat{\varphi}$ لـ φ. كما يلي:

$$\hat{g}(x)=\frac{\sum_{i=1}^{n}K\left(\frac{x-X_i}{h_n}\right)}{nh_n}$$

$$\hat{\varphi}(x)=\frac{\sum_{i=1}^{n}Y_iK\left(\frac{x-X_i}{h_n}\right)}{nh_n}$$

المقدر الدالي $\hat{f}(x)$ هو مقدر بطريقة النواة يعرف في الأخير كما يلي[1]:

$$\hat{f}(x)=\sum_{i=1}^{n}Y_iK\left(\frac{x-X_i}{h_n}\right)\Big/\sum_{i=1}^{n}K\left(\frac{x-X_i}{h_n}\right)$$

حيث h_n تسمى بالنافذة Bandwidth أو معلم التمهيد و هو العنصر المحدد لنوعية المقدر فهو يعبر عن متتالية حقيقية موجبة تؤول إلى الصفر لما حجم العينة يؤول إلى ما لا نهاية.

أما $K(.)$ تسمى بالنواة و التي تحقق الخصائص التالية:

1- و يسمى أيضا بمقدر Nadaraya-Watson.

$$\lim_{|u|\to\infty} |u|.|K(u)| = 0$$

$$\int_{-\infty}^{\infty} |K(u)| du < \infty$$

$$\sup_{u\in R} |K(u)| < \infty$$

$$\int_{-\infty}^{\infty} K(u) du = 1$$

يعتبر هذا المقدر ممهدا خطيا يعتمد على الاختيار الأمثل للنافذة h_n لأنه يحدد درجة التمهيد للمقدر حيث أثبت بعض الإحصائيين من خلال المحاكاة أنه إذا كان عدد المشاهدات المستعملة تتزايد، فإن قيمة النافذة تتناقص. إضافة إلى ذلك، كلما كبرت h_n، كلما صغر التباين ولكن مقدار التحيز يزداد. وعليه، نختار نافذة تحقق التوازن بصفة مقاربة بين التباين و مقدار التحيز[1].

إذا كبر حجم العينة، فقيمة النافذة تتناقص و يكون تقارب $\hat{f}$ نحو f بنسبة احتمال أمثلا. من أجل $\alpha \in [1/5, 1]$ لكل نقطة استمرارية $\sigma^2(x)$ و من أجل $x \in R$ حيث $g(x) > 0$. عندما $n \to \infty$، لدينا:

$$\hat{f}(x) \xrightarrow{p.s} f(x)$$

$$\sqrt{nh_n}\left(\hat{f}(x) - f(x)\right) \xrightarrow{\ell} N\left(0, \frac{\sigma^2(x)\tau^2}{g(x)}\right)$$

من أجل $\alpha \in [1/2, 1]$ لكل نقطة استمرارية $\sigma^2(x)$ و من أجـ ـل $x \in R$ حيـ ـث $2g(x) > 1-\alpha$. عندما $n \to \infty$، لدينا:

$$\sqrt{nh_n}\left(\hat{f}(x) - f(x)\right) \xrightarrow{\ell} N\left(0, \frac{\sigma^2(x)\tau^2 g(x)}{2g(x) - (1-\alpha)}\right)$$

مع $\tau^2 = \int_{-\infty}^{\infty} K^2(x) dx$.

يمكن إذن تقدير تباين القانون الطبيعي المحدود، فالتباين الشرطي $\sigma^2(x)$ المقدر بطريقة Nadaraya-Watson يعطى على الشكل التالي:

1- أنظر مثلا Chikhi (2001), p. 89

$$\hat{\sigma}^2(x) = \frac{1}{\hat{g}(x)} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right)\left(Y_i - \hat{f}(x)\right)^2$$

هناك طرق أخرى باستعمال طريقة النواة منها الانحدار كثير الحدود المحلي Local Polynomial Regression و الذي يستخدم طريقة المربعات الصغرى المبوبة Weighted Least Squares فهي طريقة مشابهة لطريقة النواة و التي من خلالها يتم تحديد الأوزان المناسبة للمعطيات[1].

في الواقع الاقتصادي، المتغير Y_i لا يفسَّر فقط بمحدد واحد و إنما قد يتأثر بمجموعة من العوامل الخارجية و لهذا ينبغي إدماج جميع العوامل المؤثرة في الظاهرة لكي يكون التحليل كاملا. في بعض الأحيان، تكون العلاقة بين الظاهرة الاقتصادية و محدداتها أو بعض محدداتها غير خطية مما يستوجب إيجاد شكل العلاقة الذي يربط هذه المتغيرات و لكن من الصعب تحديد شكلها و عليه يتم اختيار الأسلوب اللامعلمي لتعميم هذه العلاقة. نقترح في هذه الحالة نوع من النماذج وهو الشكل غير الخطي التجميعي التالي:

$$Y_i = c + f_1(X_{i1}) + f_2(X_{i2}) + \ldots + f_k(X_{ik}) + \varepsilon_i$$

نلاحظ أن Y_i مشروح من طرف k متغير مُفَسِّر و لا يمكن للمتغيرات الـ k. أن تفسر Y بشكل تام، لأنه لا يمكننا في غالب الأحيان حصر كل الظواهر المؤثرة على Y، لذلك يُدرج الخطأ العشوائي ε_i الذي يتضمن كل المعلومات التي لا تقدمها المتغيرات المفسرة.

يتم تقدير النموذج أيضا باستعمال طريقة النواة اعتمادا على أسلوب Nadaraya-Watson[2]

1- لمزيد من التفاصيل و التدقيق أكثر لهذه الطريقة، أنظر (1979) Cleveland و (1995) Mays.
2- نظرا للتعقيدات الجبرية للصيغ الرياضية، يمكن للقارئ التعمق أكثر في هذا النوع من المقدرات وذلك بالنظر إلى .
Hastie, T, Tibshirani, R and Freedman, J. و Bowman, W and Azzalini, A. (1997)
.(2001)

2.3. منهجية اختيار المعالم

في التقدير اللامعلمي، يعتبر التبويب مهم لأن طرق التقدير تعتمد عليه، حيث الأوزان تعرف رياضيا كما يلي:

$$W_i(x) = K\left(\frac{x - X_i}{h_n}\right) \Big/ \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right)$$

وهذا يعني أن مقدر Nadaraya-Watson يسمح بتقدير دالة الانحدار. السؤال الذي نطرحه: كيف يتم تحديد النافذة واختيار النواة؟

أما بالنسبة للنواة، هناك أنواع كثيرة من النواة منها النواة الطبيعية، Epanechnikov، المثلثية، المستطيلة، الثنائية،..الخ. لعل أشهر هذه الأنواع:

- نواة Epanechnikov : $K(u) = \frac{3}{4}(1-u^2)I(|u| \le 1)$
- النواة الطبيعية Gaussian Kernel: $K(u) = (2\pi)^{-1/2} \exp\left(-\frac{1}{2}u^2\right)$
- النواة المثلثية Triangular Kernel: $K(u) = (1-|u|)(I(|u| \le 1))$
- النواة المستطيلة Rectangular Kernel: $K(u) = \frac{1}{2}(I(|u| \le 1))$

يمكن القول أن النواة الطبيعية هي الأكثر استخداما في التقدير لأنها تعطي أوزانا مهمة للمتغير X_i القريب من X_n. استخدام الأنواع الأخرى من النواة يحتاج استخدام نافذة أكبر (معلم تمهيد أكبر).

يعد اختيار معلم التمهيد "النافذة" لأي مقدر غير معلمي مشكلة بالنسبة للإحصائي. تطبيقيا، يمكن الحصول على القيمة المثلى للنافذة بالتوفيق بين التباين و مقدار التحيز و يتم ذلك بصغير معيار يسمى بدالة « Cross Validation » الذي يعرف رياضيا كما يلي[1]:

$$CV(h) = \sum_{i=1}^{n} \left(Y_i - \hat{f}(X_i; h_n)\right)^2$$

1- أنظر Hârdle (1990)

مثال 6:

نختبر هذه الطريقة على متغير مصطنع (عن طريق محاكاة). عدد مشاهدات هذه السلسلة 1000. لنعتبر السلسلة غير الخطية المعرفة كما يلي:

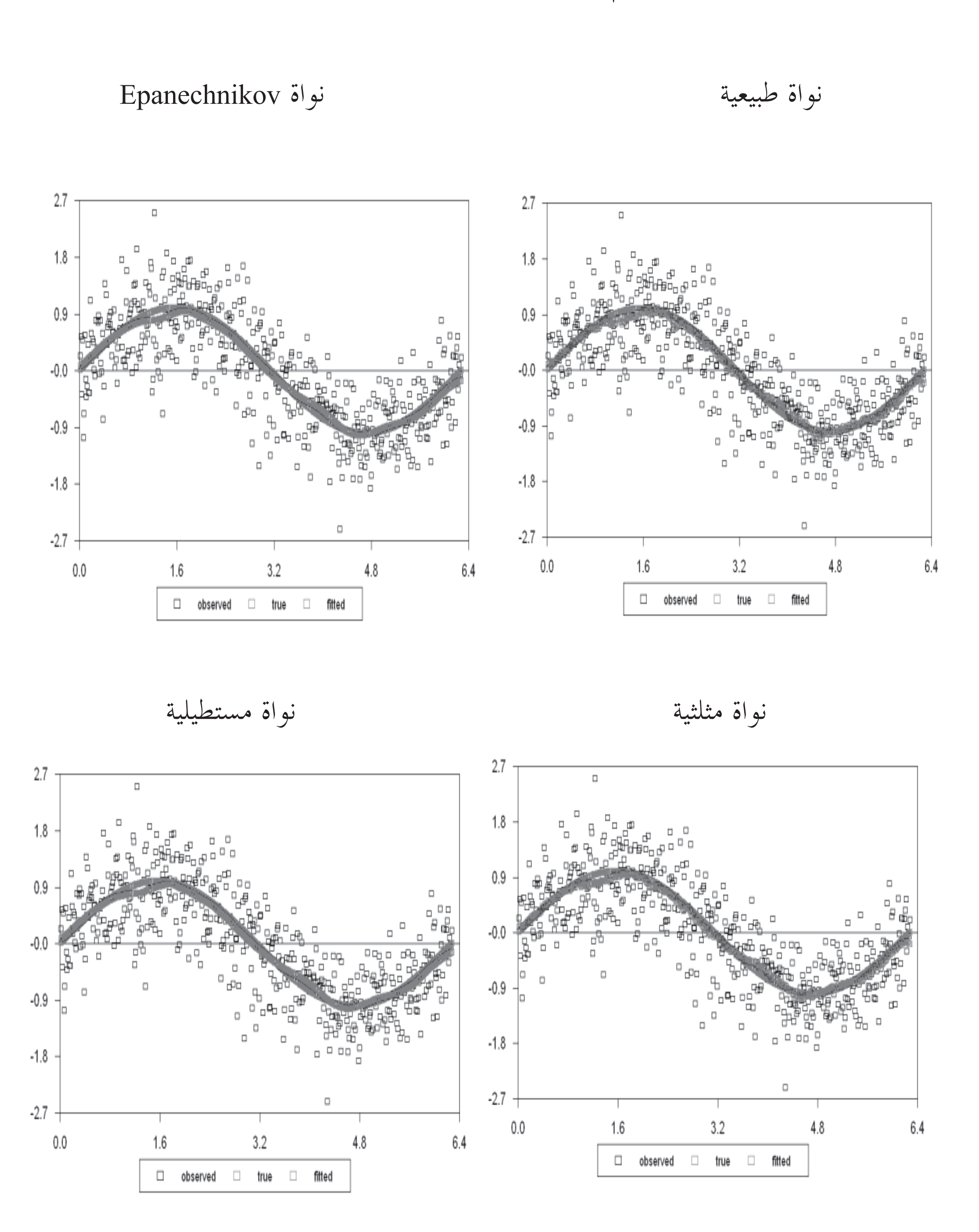
$$\omega = 2\pi i / n$$
$$Y_i = \sin\omega + \varepsilon_i, \quad \varepsilon_{\text{i}} \to N(0,\ 0.05)$$

نقوم بحساب قيم مختلفة للنافذة h_n بطريقة Cross validation. ترتكز إذن هذه الطريقة على حساب النافذة و التي تصغر معيار Cross validation من أجل كل نواة. النتائج تظهر في الجدول التالي:

الجدول (4): تقدير النافذة و دالة Cross Validation بطريقة النواة

شكل النواة Kernel form							
المستطيلي Rectangular		Triangular المثلثي		Epanechnikov		Gaussian الطبيعي	
CV	h_n	*CV*	h_n	*CV*	h_n	*CV*	h_n
0.2946	0.005	0.3104	0.005	0.2703	0.005	0.2792	0.005
0.2759	0.010	0.2836	0.010	0.2619	0.010	0.2669	0.010
0.2587	**0.05**	0.2603	0.05	0.2658	0.05	0.2596	0.05
0.2588	0.075	0.2589	0.075	0.2882	0.075	0.2654	0.075
0.2683	0.015	0.2742	0.015	0.2590	0.015	0.2625	0.015
0.2626	0.020	0.2690	0.020	**0.2585**	**0.020**	0.2605	0.020
0.2601	0.035	0.2626	0.035	0.2599	0.035	**0.258**	**0.035**
0.2594	0.08	**0.2589**	**0.08**	0.2945	0.08	0.2672	0.08
0.2589	0.065	0.2592	0.065	0.2700	0.065	0.2624	0.065

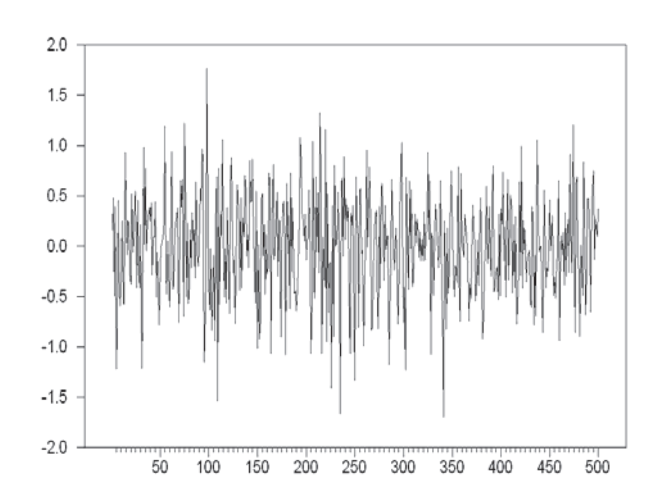
الشكل رقم (2): تقدير الانحدار بطريقة النواة



قمنا بتقدير الانحدار غير المعلمي لـ Y على ω بطريقة Nadaraya-Watson الذي يرتكز أساسا على التمهيد باستعمال النواة معتمدا على تقدير التوقع الرياضي الشرطي. العنصر المحدد لجودة المقدر هو الاختيار الأمثل للنافذة. يظهر جليا من خلال الشكل الموضح أعلاه أن القيم الحقيقية و الفعلية تنطبق تقريبا أي تتقارب فيما بينها كما يبدو لنا أيضا أن كل أنواع النواة المستخدمة هنا تعطي نفس النتائج.

نلاحظ أن للنموذج (باستعمال النواة الطبيعية) قدرة تفسيرية $R^2 = 0.6425$ و بواقي التقدير المبينة في الشكل (3) مستقلة ذاتيا حيث إحصائية دربين–واتسون تساوي 2.03.

الشكل رقم (3): بواقي التقدير (النواة الطبيعية) عند $h_n = 0.035$



نقترح الآن النموذج غير الخطي التجميعي التالي:

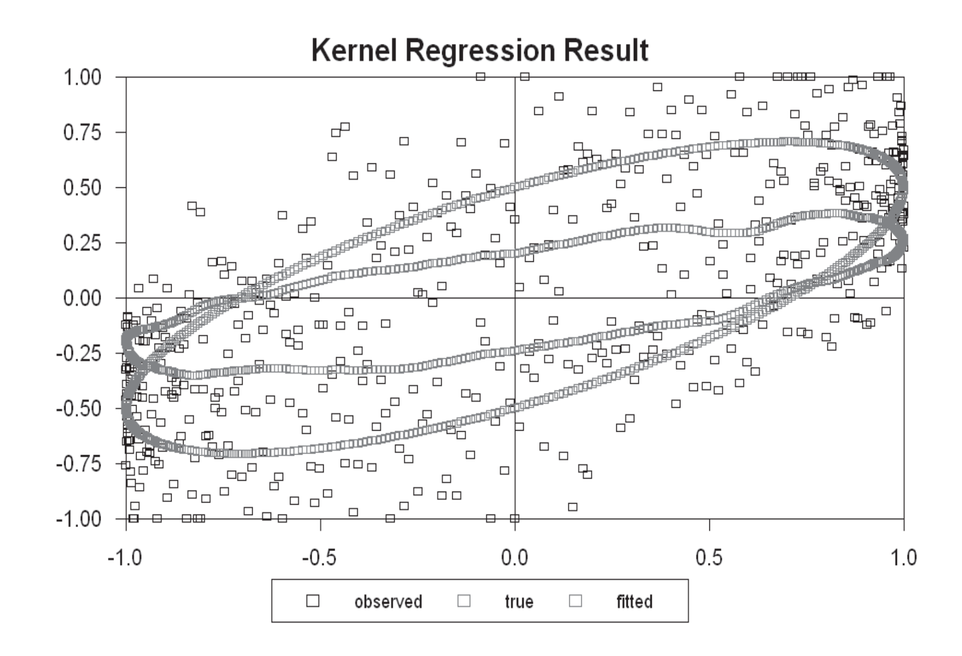
$$Y_i = f_1(X_{i1}) + f_2(X_{i2}) + \varepsilon_i$$

حيث:

$$X_{i1} = \sin\omega$$
$$X_{i2} = \cos\omega$$

لتقدير الانحدار المتعدد لمتغيرين، يكفي تقدير Y_i على X_{i1}, X_{i2} باستخدام النواة المثلثية.

الشكل رقم (4): تقدير الانحدار المتعدد بطريقة النواة المثلثية



حيث تم تقدير النموذج باستعمال نافذة تساوي 0.10. يمكن القول أن الاختيار الأمثل للنافذة يضمن سرعة تقارب مثلى. لتقدير هذا النوع من النماذج استعنا ببرنامج RATS 5.04.

لاختيار النافذة المثلى، نقوم بتصغير معيار Cross validation:

```
declare vector[series] yHat_vec(7)
compute [vector] bandwidth_vec = ||0.005,0.01,0.05,0.075,0.015,0.02,0.035,0.08,0.065||
do ii=1,9
 @kernreg(kernel=gaussian,bandwidth=relative) y / yHat_vec(ii) eps
 # w
 # bandwidth_vec(ii)
 display "relative bandwidth:" #.### bandwidth_vec(ii) $
         ";  cross-valuation function:" #.###### %cv
end do
scatter(header="Kernel Regression Results",$
        subheader="with different bandwidths",$
        style=dotline,key=below,klabel=||"observed","h=.5%","h=1%","h=5%"||) 4
# w y           / 1
# w yHat_vec(1)  / 4
# w yHat_vec(2)  / 6
# w yHat_vec(3)  / 2
```

نقدر الانحدار على النافذة المثلى، مثلا باستعمال النواة الطبيعية:

```
@kernreg(kernel=gaussian,bandwidth=relative) y / yHat resids
 # w
 # 0.035
```

أما بالنسبة للانحدار المتعدد:

```
@kernreg(kernel=triangular,bandwidth=relative) z / zHat eps
# xTrue  yTrue
# 0.10   0.10
scatter(header="Kernel Regression Result",style=dots,$
        hmin=-1.0,hmax=1.0,vmin=-1.0,vmax=1.0,$
        key=below,klabel=||"observed","true","fitted"||) 3
# xTrue z     / 1
# xTrue zTrue / 2
# xTrue zHat  / 4
```

الفصل الخامس

مدخل إلى نماذج المعادلات الآنية

الفصل الخامس

مدخل إلى نماذج المعادلات الآنية

في الفصول السابقة، كان التقدير مقتصرا على النماذج الاقتصادية المكونة من معادلة واحدة تحوي متغيرا تابعا واحدا، وقد تحوي العديد من المتغيرات المستقلة. في مجال الاقتصاد يوجد نماذج تتألف من عدة معادلات يجمع بينها تأثير مشترك بواسطة المتغيرات المتضمنة في النموذج. معظم تطبيقات الاقتصاد تتكون من العديد من المعادلات التي تنتمي إلي نظام متداخل. من أهم النماذج في الاقتصاد نموذج العرض والطلب حيث يتم تداخل تحديد السعر والكميه بين النموذجين. فمثلا لدراسة الطلب على سلعة ما يجب دراسة العرض نظرا لتداخل النموذجين معا. في المعادلات الآنية، لا يمكن تطبيق طريقة المربعات الصغرى العادية لأن المتغيرات المستقلة قد تتضمن متغيرات تابعه ويوجد ترابط بين المتغيرات والخطأ العشوائي مما يؤدي إلى الحصول على مقدرات متحيزة. لذلك يتم استخدام طرق أخرى للتقدير منها طريقة المربعات الصغرى على مرحلتين أو المربعات الصغرى غير المباشرة بدلا من المربعات الصغرى العادية.

في هذا الفصل، نقوم بدراسة المعادلات الآنية و ذلك باقتراح طريقة لتقدير معالم الشكل الهيكلي و دراسة الخصائص الإحصائية للمقدرات و بعد ذلك سنتطرق إلى مشكل التمييز و اختبار الآنية.

1. أمثلة على نماذج المعادلات الآنية:

1.1. نموذج العرض والطلب:

كما هو معروف فان سعر السلعة والكمية المباعة تتحدد عن طريق التفاعل بين منحنى العرض والطلب للسلعة. للتبسيط، نفترض أن منحنيات العرض والطلب خطيه وبإضافة المتغير العشوائي يمكن كتابة المعادلة كما يلي:

دالة الطلب: $\alpha < 0, \quad Q_{dt} = \alpha_0 + \alpha_1 P_t + \varepsilon_{1t}$

دالة العرض: $\beta > 0 \,, \quad Q_{st} = \beta_0 + \beta_1 P_t + \varepsilon_{st}$

التوازن: $Q_{dt} = Q_{st}$

حيث:

Q_{dt}: الكميه المطلوبة في الفترة الزمنية t.

Q_{st}: الكميه المعروضة في الفترة الزمنية t.

P_t : سعر السلعة في الفترة الزمنية t.

ε_t : الخطأ العشوائي.

نلاحظ أن ل P . و Q تأثير متبادل فمثلا المتغير العشوائي ε_{1t} يتغير بسبب التغير في المتغيرات التي تؤثر على Q مثل الدخل، الثروة و الذوق، فينتقل منحنى الطلب إلى أعلى إذا كانت موجبة والى اليسار إذا كانت ε_{1t} سالبة، أي انتقال المنحنى يؤدي إلى تغير قيمة P و Q. وكذلك إذا تغيرت ε_{2t} (تغير سعر عناصر الإنتاج، تغيرات في التقنية...الخ) ستؤدي إلى انتقال منحنى العرض مسببة في تغير P و Q بسبب الارتباط المتداخل بين P، Q و ε_{1t}. في المعادلة الأولى P و Q و ε_{2t} و P في المعادلة الثانية. لذلك لا يمكن تطبيق المربعات الصغرى العادية بسبب عدم تحقق فرضيات الاستقلالية بين المتغيرات المفسرة والمتغير العشوائي.

2.1. نموذج السعر و الأجور:

ليكن نموذج فيليب للأجور والنقود والسعر المعرف رياضيا كما يلي:

$$W_t = \alpha_0 + \alpha_1 UN_t + \alpha_2 P_t + \varepsilon_{1t}$$
$$P_t = \beta_0 + \beta_1 W_t + \beta_2 R_t + \beta_3 M + \varepsilon_{2t}$$

W_t: معدل التغير في الأجور

UN_t: معدل البطالة

P_t : معدل التغير في الأسعار

R_t : معدل التغير في تكلفة رأس المال

M_t : معدل التغير في الأسعار لعناصر الإنتاج المستوردة

t: الزمن

ε_{1t} و ε_{2t} هي المتغيرات العشوائية.

حيث أن المتغير P يدخل في معادلة الأجور والمتغير W يدخل في معادلة السعر، مما يعني أن المتغيرين ثنائي التأثير. بناء على ذلك تكون المتغيرات المستقلة مرتبطة مع المتغيرات العشوائية مما يؤدي إلى عدم تحقق الفرضيات الخاصة بطريقة المربعات الصغرى العادية و لا يمكن تطبيقها لتقدير النموذج.

3.1. نموذج كينيز لتحديد الدخل:

لنعتبر النموذج الاقتصادي الكلي المكون من ثلاث معادلات:

$$C_t = a_0 + a_1 Y_t + \varepsilon_{1t} \;,\quad 0 < a_1 < 1$$

$$I_t = b_0 + b_1 Y_{t-1} + \varepsilon_{2t}$$

$$Y_t = C_t + I_t$$

C_t: الاستهلاك الإجمالي في الفترة t

I_t: الاستثمار الإجمالي في الفترة t

Y_t: الدخل الوطني في الفترة t

ε_{1t} و ε_{2t} هي المتغيرات العشوائية.

نلاحظ أن الدخل يفسر الاستهلاك أي يظهر كمتغير مستقل في المعادلة الأولى أما في المعادلة الثالثة يظهر كمتغير تابع و هذا يعني أن المتغيرين ثنائي التأثير. من جهة أخرى، هناك استقلالية تامة بين Y_t و ε_{1t} في معادلة الاستهلاك و هذا ما يتناقض مع معادلة التوازن، حيث أنه إذا قمنا بتعويض C_t و I_t بما يساويهما في المعادلة الثالثة، نحصل على ما يلي:

$$Y_t = a_0 + a_1 Y_t + \varepsilon_{1t} + b_0 + b_1 Y_{t-1} + \varepsilon_{2t}$$

أي:

$$(1 - a_1) Y_t = a_0 + b_0 + b_1 Y_{t-1} + \varepsilon_{2t} + \varepsilon_{1t}$$

و عليه:

$$Y_t = \frac{a_0 + b_0}{1 - a_1} + \frac{b_1}{1 - a_1} Y_{t-1} + \frac{\varepsilon_{2t} + \varepsilon_{1t}}{1 - a_1}$$

و هذه الأخيرة تشير إلى أن المتغير Y_t دالة تابعة لـ ε_{1t} . أي $\operatorname{cov}(Y_t, \varepsilon_{1t}) \neq 0$ وهذا تناقض و خرق لفرضيات المربعات الصغرى العادية، مما يؤدي إلى الحصول على مقدرات متحيزة و غير متسقة.

2. البناء الهيكلي والصورة المختزلة للمعادلات:

لمعرفة المشاكل التي تواجه تقدير المعادلات الآنية، يجب تعريف بعض المفاهيم مثل طبيعة نظام المعادلات الآنية، هل التغير في السعر مثلا هو الذي يسبب ارتفاع الكمية المطلوبة أم ارتفاع الكمية المطلوبة هو الذي يؤدي إلى التغير في السعر؟ هناك تحديد مرتبط بين المتغيرين أي أن هناك علاقة سببية بين المتغيرين. ارتفاع الأسعار سيؤدي إلى انخفاض الكمية المطلوبة، وانخفاض الكمية المطلوبة سيؤدي إلى انخفاض الكمية المعروضة وانخفاض الكمية المعروضة سيؤدي إلى ارتفاع الأسعار وهكذا. إن الاقتصاد مليء بالكثير من الأمثلة بالتأثير المرجعي والسببية الثنائية مما يتطلب تطبيق المعادلات الآنية.

يسمى نظام المعادلات المتعددة، المقترح من طرف الاقتصادي و الذي يترجم مباشرة العلاقات بين المتغيرات بنظام المعادلات الهيكلية و نظام المعادلات الآنية هو النظام الذي يكون هناك تأثير لـ Y على ألأقل على أحد المتغيرات المستقلة بالإضافة إلى التأثير الموجود من المتغيرات المفسرة على المتغير التابع. لبناء نموذج المعادلات الآنية، يجب الفصل بين المتغيرات التي تحدد آنياً (Y_{1t} و Y_{2t} تسمى المتغيرات الداخلية) و المتغيرات التي لا تحدد آنياً (X_{1t} و X_{2t} وتسمى متغيرات خارجية).

$$Y_{1t} = \alpha_0 + \alpha_1 Y_{2t} + \alpha_2 X_{1t} + \alpha_3 X_{2t} + \varepsilon_{1t}$$
$$Y_{2t} = \beta_0 + \beta_1 Y_{1t} + \beta_2 X_{1t} + \beta_3 X_{2t} + \varepsilon_{2t}$$

إذا كانت مثلا Y_1 هي الكمية المطلوبة للحم الخروف و Y_2 سعر الخروف و X_1 دخل المستهلكين و X_2 سعر لحم البقر (السلعة البديلة). أي أن المعادلة الأولى تمثل سلوك المستهلك بينما المعادلة الثانية تمثل سلوك المنتج والمعادلتان معا تسمى بنظام *المعادلات الهيكلية*.

تصف المعادلات الهيكلية النظرية الاقتصادية خلـ ف المـ تغيرات الداخلية بـ التعبير بالمصطلحين الداخلي والخارجي. يجب أن ينظر الباحثون إلى النظام كاملاً ليتمكنوا مـ ن معرفة المردود الالتفافي المتداخل في النظام. على سبيل المثـ ال $Y_{m,t}$ محـ ددة بالاشـ تراك (jointly determined). أي تغير في Y_{1t} سيؤدي إلى تغير في Y_{2t} والتي تؤدي بـ دورها إلى تغير في Y_{1t}. مقارنة بالتغير في X_{1t} والذي سيؤدي إلى التغير في Y_{1t} ولكن لن يلتف مرتدا ليؤثر في X_{1t} مره أخرى. المعـ املات $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_0, \beta_1, \beta_2, \beta_3$ تسـ مى المعاملات الهيكلية واختبار الفرضية يجب أن يكون عن قيمهم ومؤشـ راتهم (سـ الب أو موجب) كما هو في معاملات المعادلة الواحدة.

نلاحظ أن المتغير يسمى داخلي لأنه مشترك التحديد وليس لأنـ ه يظهـ ر في كـ لا المعادلتين. أي أن X_{2t} وهو سعر لحم البقر يعتبر متغيرا خارجيا لأنه لا يتحـ دد آنيـ اً في سوق الخراف. السؤال المطروح هنا: كيف نفرر ما إذا كان متغيرا داخليا أو خارجيـ ا؟ هناك بعض العوامل دائماً خارجية مثل الجو الخارجي، ولكن هناك متغيرات لا تعتـ بر لا خارجية ولا داخلية اعتمادا على عدد وطبيعة المتغيرات في المعادلات الأخرى في النظـ ام، ففي بعض النماذج التي تتضمن متباطئات 0 نظام يتضمن معادلات المتباطئـ ات الموزعـ ة وللتوضيح تسمى المتغيرات المتباطئة للمتغير الداخلي والمتغيرات الخارجيـ ة في النمـ وذج بالمتغيرات المحددة سابقاً، أي أن المتغيرات الخارجية والمتغيرات المتباطئة محدده خارج النظام لمعادلات محدده أو قبل الفترة الحالية.

إن الصورة المختزلة للمعادلات الآنية تعبر عن كتابة كل متغير داخلي بدلالـ ة جميـ ع المتغيرات الخارجية في النموذج الآني، من خلال النموذج الكيتري، لدينا:

$$Y_t = \frac{a_0 + b_0}{1 - a_1} + \frac{b_1}{1 - a_1} Y_{t-1} + \frac{\varepsilon_{2t} + \varepsilon_{1t}}{1 - a_1}$$

$$I_t = b_0 + b_1 Y_{t-1} + \varepsilon_{2t}$$

$$C_t = \frac{a_0 + a_1 b_0}{1 - a_1} + \frac{a_1 b_1}{1 - a_1} Y_{t-1} + \frac{a_1 \varepsilon_{2t} + \varepsilon_{1t}}{1 - a_1}$$

حيث تم كتابة كل من C_t، I_t و Y_t بدلالة المتغير الخارجي (المحدد مسبقا) و هو Y_{t-1}.

3. الصيغة العامة لنماذج المعادلات الآنية و التحيز الآني:

لتعميم الصيغ السابقة لنماذج المعادلات الآنية، نفترض أنه لدينا الهيكل العام لنموذج خطي متكون من m معادلة هيكلية و كل معادلة تحتوي على m متغير داخلي و k متغير محدد مسبقا و متغيرات الحد العشوائي الموزعة توزيعا طبيعيا. يمكن كتابة النموذج رياضيا كما يلي:

$$b_{11}Y_{1t} + b_{12}Y_{2t} + \ldots + b_{1m}Y_{mt} + c_{11}X_{1t} + c_{12}X_{2t} + \ldots + c_{1k}X_{kt} = \varepsilon_{1t}$$
$$b_{21}Y_{1t} + b_{22}Y_{2t} + \ldots + b_{2m}Y_{mt} + c_{21}X_{1t} + c_{22}X_{2t} + \ldots + c_{2k}X_{kt} = \varepsilon_{2t}$$
$$\ldots\ldots$$
$$b_{m1}Y_{1t} + b_{m2}Y_{mt} + \ldots + b_{mm}Y_{mt} + c_{m1}X_{1t} + c_{m2}X_{2t} + \ldots + c_{mk}X_{kt} = \varepsilon_{mt}$$

و بشكله المصفوفي:

$$\begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1m} \\ b_{21} & b_{22} & \ldots & b_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ b_{m1} & b_{m2} & \ldots & b_{mm} \end{pmatrix} \begin{pmatrix} Y_{1t} \\ Y_{2t} \\ \ldots \\ Y_{mt} \end{pmatrix} + \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1k} \\ c_{21} & c_{22} & \ldots & c_{2k} \\ \ldots & \ldots & \ldots & \ldots \\ c_{m1} & c_{m2} & \ldots & c_{mk} \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \\ \ldots \\ X_{kt} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \ldots \\ \varepsilon_{mt} \end{pmatrix}$$

أي:

$$\underset{(m,m)}{B} . \underset{(m,1)}{Y} + \underset{(m,k)}{C} . \underset{(k,1)}{X} = \underset{(m,1)}{\varepsilon}$$

بطبيعة الحال، في كل معادلة، هناك بعض المعاملات معدومة و المتغير الذي معامله يساوي الواحد يعتبر المتغير التابع. إذا كانت المصفوفة B معرفة، فيمكن الانتقال من الشكل الهيكلي إلى الشكل المختزل و ذلك بكتابة الشعاع Y بدلالة الشعاع X حيث:

$$Y = -B^{-1}CX + B^{-1}\varepsilon$$

فيمكن إذن تطبيق طريقة المربعات الصغرى العادية باعتبار أن الأخطاء $B^{-1}\varepsilon$ مستقلة عن X.

بالرغم من بساطة الصيغة الرياضية إلا أن جانبه التطبيقي معقد نوعا ما، فمعرفة $m \times k$ عنصر للمصفوفة $B^{-1}C$ لا يسمح بتحديد المصفوفة B التي تحتوي على $m \times m$

عنصر و هذا فضلا عن المصفوفة C المكونة من $m\times k$ عنصر. نحن إذن في حالة وجود $m\times k$ معادلة لـ $(m\times m)+(m\times k)$ مجهول، فبدون قيود إضافية، من المستحيل إيجاد حلول ممكنة. الأمر هنا يتعلق بمشكل التمييز (أو التعريف) Identification.

بالعودة إلى النموذج الكيتري، يمكن كتابة الشكل المصفوفي حيث:

$$Y=\begin{pmatrix} C_t \\ I_t \\ Y_t \end{pmatrix} \quad ، \quad B=\begin{pmatrix} 1 & 0 & -a_1 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \quad ، \quad X=\begin{pmatrix} 1 \\ Y_{t-1} \end{pmatrix} \quad ، \quad C=\begin{pmatrix} -a_0 & 0 \\ -b_0 & -b_1 \\ 0 & 0 \end{pmatrix} \quad و$$

$$\varepsilon=\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ 0 \end{pmatrix}$$

لاستخدام طريقة المربعات الصغرى يجب أن تتحقق جميع الفرضيات الأساسية لأنه إذا تم تطبيق هذه الطريقة على المعادلات الهيكلية للنظام الآني فإن المقدرات تكون متحيزة. هذا التحيز يسمى التحيز الآني أو تحيز المعادلات الآنية، بمعنى أنه في النظام الآني، القيم المتوقعة لمقدرات المربعات الصغرى للمعاملات الهيكلية لا تساوي القيمة الحقيقية $E(\hat{\beta})\neq\beta$. يحدث هذا التحيز عندما يرتبط الخطأ العشوائي خطيا مع المتغيرات الداخلية في النموذج (المتغير التابع يدخل في المعادلة كمتغير مفسر).

4. مشكل التمييز (التعريف) The identification Problem:

تشير مشكلة التمييز إلى إمكانية أو عدم إمكانية حساب المعالم الهيكلية لنموذج المعادلات الآنية انطلاقا من معالم النموذج المختزل. يجب دراسة مشكلة التمييز حيث أنه لا يمكن تطبيق طريقة تقدير مناسبة على المعادلات إلا إذا كانت هذه الأخيرة معرفة، فإذا كانت تلك المعادلات معرفة، فإنه يمكن تقدير معالم الشكل الهيكلي. أما إذا كانت المعادلة غير مميزه فإن ذلك يعني أنه لا يمكن تقدير المعالم الهيكلية لنماذج المعادلات الهيكلية انطلاقا من معالم الصيغة المختزلة.

لدراسة شروط التمييز (التعريف)، لابد من أن نشير إلى أن هناك قيود على المعاملات وهما نوعان: قيود الإقصاء و القيود الخطية. أما بالنسبة لقيود الإقصاء، نعتبر كل مرة أن متغيرا داخليا أو خارجيا لا يظهر في المعادلة الهيكلية و هذا يرجع لكون أن هذه المتغيرات لها معامل معدوم. على سبيل المثال، إذا رجعنا إلى النموذج الكيتري، نلاحظ أن المتغير I_t غير موجود في معادلة الاستهلاك، معامله يساوي إذن الصفر. في المصفوفة B، عنصر السطر الأول و العمود الثاني يساوي الصفر، أما القيود الخطية يتعلق الأمر هنا بوجود قيود على المعالم حيث أن بعض المتغيرات قد تشترك في معامل واحد و هذا ما نلاحظه في بعض النماذج الاقتصادية..

تتحدد شروط التعريف معادلة بمعادلة. هناك ثلاث حالات للتعريف:

- المعادلة ناقصة تعريف under-identified، إذا كان عدد المتغيرات الخارجية في المعادلة يتجاوز عدد المتغيرات الداخلية في المعادلة مطروحا منه واحد. الحل في النظام الهيكلي مستحيل.
- المعادلة معرفة تماما exactly-identified، إذا كان عدد المتغيرات الخارجية في العادلة مساويا لعدد المتغيرات الداخلية في المعادلة مطروحا منه واحد.
- المعادلة زائدة تعريف over-identified، إذا كانت عدد المتغيرات الخارجية في المعادلة يقل عن عدد المتغيرات الداخلية في المعادلة مطروحا منه واحد.

إذا كان النموذج ناقص تعريف، فإن ليس هناك إمكانية لتقدير معالم النموذج و بالتالي ينبغي إعادة النمذجة.

تطبيقيا، هناك قاعدة سهلة لدراسة شروط التمييز، لدينا أولا:

m : عدد المتغيرات الداخلية في النموذج (أو أيضا عدد المعادلات)

k : عدد المتغيرات الخارجية في النموذج

m' : عدد المتغيرات الداخلية التي تظهر في معادلة ما

k' : عدد المتغيرات الخارجية التي تظهر في معادلة ما

عندما تكون القيود إلا قيود الإقصاء، الشروط الضرورية للتمييز هي كالتالي:

- ❖ $m-m^{'}+k-k^{'}<m-1$: المعادلة ناقصة تعريف
- ❖ $m-m^{'}+k-k^{'}=m-1$: المعادلة معرفة تماما
- ❖ $m-m^{'}+k-k^{'}>m-1$: المعادلة زائدة تعريف

عندما يكون لدينا r قيد يتعلق الأمر بالقيود على المعالم، الشروط تصبح كما يلي:

- ❖ $m-m^{'}+k-k^{'}+r<m-1$: المعادلة ناقصة تعريف
- ❖ $m-m^{'}+k-k^{'}+r=m-1$: المعادلة معرفة تماما
- ❖ $m-m^{'}+k-k^{'}+r>m-1$: المعادلة زائدة تعريف

و هذه الشروط هي ضرورية و ليست كافية و تسمى أيضا شروط الترتيب. يجب إذن على الإحصائي التحقق من الشروط الكافية و التي تسمى بشروط الرتبة و التي تعتبر صعبة تطبيقيا.

ليكن النموذج على الشكل المصفوفي:

$$\underset{(m,m)}{B}\,.\underset{(m,1)}{Y}+\underset{(m,k)}{C}\,.\underset{(k,1)}{X}=\underset{(m,1)}{\varepsilon}$$

لتكن المصفوفة P حيث $P=[BC]$ و التي ذات بعد $(m,m+k)$

لتكن ϕ_i مصفوفة القيود المتعلقة بالمعادلة i حيث: $P_i\phi_{ih}=0$ ، i هو السطر رقم i للمصفوفة P و h العمود رقم h للمصفوفة ϕ_i.

لتكن $\mu_i=rang[P\phi_i]$ رتبة المصفوفة $P\phi_i$ و m عدد المتغيرات الداخلية في النموذج، شرط الرتبة هو كالتالي:

- ❖ $\mu_i<m-1$: المعادلة i ناقصة تعريف
- ❖ $\mu_i=m-1$: المعادلة i معرفة تماما
- ❖ $\mu_i>m-1$: المعادلة i زائدة تعريف

مثال 1[1]:

ليكن النموذج التالي:

$$Y_{1t} = a(Y_{2t} + X_t) + \varepsilon_{1t}$$
$$Y_{2t} = bY_{1t} + cY_{1,t-1} + \varepsilon_{2t}$$

حيث:

Y_{1t}: الناتج الوطني الإجمالي خلال السنة t

Y_{2t}: استهلاك الأسر خلال السنة t

X_t: الطلب النهائي خلال السنة t

المطلوب تحديد شروط تمييز النموذج (الضروري و الكافي)

يحتوي النموذج على متغيرين داخليين Y_{1t}، Y_{2t} أي $m = 2$ و متغيرين خارجيين X_t و Y_{1t-1} أي $k = 2$. نلاحظ أن المعادلة الأولى تحتوي على قيد خطي و الذي يعبر عن القيد حول المعاملات حيث هناك متغيران يشتركان في معلم واحد. بتطبيق شروط التمييز، نستنتج أن هناك متغيرين داخليين $m' = 2$ و متغير خارجي واحد $k' = 1$ و هناك قيد واحد حول المعالم $r = 1$، أما المعادلة الثانية، هناك متغيران داخليان و متغير خارجي واحد $m' = 2$ و $k' = 1$ إلا أنه لا يوجد قيد حول المعالم $r = 0$.

بالنسبة للمعادلة الأولى، لدينا:

$m - m' + k - k' + r = 2 - 2 + 2 - 1 + 1 = 2 > m - 1 = 2 - 1 = 1$ و عليه المعادلة الأولى زائدة تعريف.

بالنسبة للمعادلة الثانية،

$m - m' + k - k' + r = 2 - 2 + 2 - 1 + 0 = 1 = m - 1 = 2 - 1 = 1$ إذن المعادلة الثانية معرفة تماما.

1- Bourbonnais (2003), p. 212

يتم تطبيق شروط الرتبة، النموذج المصفوفي يكتب كما يلي:

$$\begin{pmatrix} 1 & -a \\ -b & 1 \end{pmatrix}\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} + \begin{pmatrix} -a & 0 \\ 0 & -c \end{pmatrix}\begin{pmatrix} X_t \\ Y_{1,t-1} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

$$B \quad \times \quad Y \quad + \quad C \quad \times \quad X \quad = \quad \varepsilon$$

لتكن المصفوفة P حيث:

$$P = [BC] = \begin{pmatrix} 1 & -a & -a & 0 \\ -b & 1 & 0 & c \end{pmatrix}$$

المصفوفة ϕ_1، مصفوفة القيود للمعادلة الأولى، تحدد كما يلي:

- كل سطر يمثل بمتغير (السطر 1 = Y_{1t}، السطر 2 = Y_{2t}، السطر 3 = X_t، السطر 4 = $Y_{1,t-1}$)
- عمود بقيد إقصاء (متغير داخلي أو خارجي غائب) و بقيد خطي على المعاملات
- بالنسبة لعلاقات الإقصاء، الأعمدة مكونة من العدد 0، باستثناء المتغيرات التي معاملاتها معدومة، نضع القيمة 1
- بالنسبة للقيود على المعاملات، نوضح العلاقة بين المعاملات. في هذا النموذج، لدينا: $c_{11} - c_{12} = 0$ (لـ Y_{2t} و X_t نفس المعامل)

يمكن إعادة كتابة الشرط: $P_i \phi_{ih} = 0$

ليس لدينا في المعادلة الأولى إلا متغير غائب واحد وقيد على معامل المتغيرين Y_{2t} و X_t، ليكن العمودان:

Y_{2t} و X_t لهما نفس المعامل

المتغير $Y_{1,t-1}$ غائب في المعادلة الأولى ←

$$\phi_i = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & -1 \\ 1 & \leftarrow 0 \end{pmatrix}$$

بالنسبة لـ ϕ_2، لدينا قيد واحد: المتغير X_t غائب في المعادلة الثانية، ليكن:

المتغير X_t غائب في المعادلة الثانية $\leftarrow \quad \phi_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$

$$P\phi_1 = \begin{pmatrix} 1 & -a & -a & 0 \\ -b & 1 & 0 & -c \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -c & 1 \end{pmatrix}$$

و بناء على ذلك، فإن: $rang(P\phi_1) = \mu_1 = m-1$

و هذا يعني أن المعادلة الأولى معرفة تماما.

$$P\phi_2 = \begin{pmatrix} 1 & -a & -a & 0 \\ -b & 1 & 0 & -c \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -a \\ 0 \end{pmatrix}$$

أي: $rang(P\phi_2) = \mu_2 = 1 = m-1$

إذن المعادلة الثانية معرفة تماما (إلا إذا كان $a = 0$)

نلاحظ أن في شرط الرتبة المعادلة الأولى معرفة تماما عكس نتيجة الشرط الضروري حيث وجدنا أن هذه المعادلة زائدة تعريف و لكن لا يهم طالما المعادلة معرفة في كلا الحالتين.

5. طرق تقدير المعادلات الآنية

إن هناك العديد من طرق التقدير التي يمكن استعمالها لتجنب التحيز الموجود في حالة تطبيق المربعات الصغرى العادية على المعادلات الآنية إلا أن أكثر طريقة مستخدمه هي طريقة المربعات الصغرى غير المباشرة (ILS) وطريقة المربعات الصغرى على مرحلتين (2SLS) و تسمى أيضا بطريقة المربعات الصغرى المضاعفة (DLS).

- إذا كان النموذج ناقص تعريف، فالتقدير مستحيل

- إذا كانت المعادلة معرفة تماما، يتم استعمال طريقة المربعات الصغرى غير المباشرة (ILS) أو طريقة المربعات الصغرى على مرحلتين (2SLS).
- أما إذا كانت العادلة زائدة تعريف، ففي هذه الحالة لا يتم تطبيق إلا طريقة المربعات الصغرى على مرحلتين (2SLS).

1.5. طريقة المربعات الصغرى غير المباشرة

ترتكز طريقة المربعات الصغرى غير المباشرة على تطبيق طريقة المربعات الصغرى العادية على المعادلات المعرفة تماما للنموذج في شكله المختزل. نتبع الخطوات التالية:

- الانتقال من الشكل الهيكلي إلى الشكل المختزل للنموذج، أي كتابة كل متغير داخلي بدلالة جميع المتغيرات الخارجية في النموذج
- تقدير كل معادلة بطريقة المربعات الصغرى العادية (OLS)
- حساب معاملات المعادلات الهيكلية عن طريق العلاقة الجبرية الموجودة بين المعاملات المختزلة و الهيكلية (الحل وحيد لأن النموذج معرف تماما)

تعتبر مقدرات طريقة المربعات الصغرى غير المباشرة (ILS) للشكل المختزل أحسن تقدير خطي و غير متحيز BLUE. غير أن مقدر معاملات الشكل الهيكلي، المتحصل عليه انطلاقا من المقدر الأمثل بطريقة المربعات الصغرى غير المباشرة (ILS)، متحيز في العينات الصغيرة، فالخصائص التقاربية تجعل من التحيز يقترب من الصفر كلما كبر حجم العينة.

2.5. طريقة المربعات الصغرى على مرحلتين أو المضاعفة

إن هذه الطريقة هي الأكثر استخداما في المجال التطبيقي، حيث أن مقدرات المربعات الصغرى العادية سوف تكون متحيزة لتجنب هذا التحيز يمكن إيجاد متغير يتميز بكونه مساويا في القيمة للمتغير الداخلي و ألا يكون مرتبطا مع الخطأ العشوائي. إذا وجد هذا المتغير وتم استبداله مع المتغير الداخلي حيث يظهر كمتغير مفسر ويكون غير مرتبط مع الخطأ العشوائي، فإن الفرضيات الأساسية للنموذج تكون محققة. يسمى هذا المتغير بالمتغير الأداة Instrumental variable ليحل محل المتغير الداخلي، حيث أنه لا توجد

سببية causality بين المتغير الأداتي وأي من المتغيرات الداخلية فاستخدام المتغير الأداتي يجنب النموذج مشكلة عدم تحقق فرضيات OLS. لإيجاد ذلك المتغير نسعى لاستخدام طريقة المربعات الصغرى على مرحلتين (أو المضاعفة).

ليكن النموذج المتكون من m متغير داخلي و k متغير خارجي و متغيرات الحد العشوائي الموزعة توزيعا طبيعيا:

$$
\begin{aligned}
& b_{11}Y_{1t} + b_{12}Y_{2t} + ... + b_{1m}Y_{mt} + c_{11}X_{1t} + c_{12}X_{2t} + ... + c_{1k}X_{kt} = \varepsilon_{1t} \\
& b_{21}Y_{1t} + b_{22}Y_{2t} + ... + b_{2m}Y_{mt} + c_{21}X_{1t} + c_{22}X_{2t} + ... + c_{2k}X_{kt} = \varepsilon_{2t} \\
& ..... \\
& b_{m1}Y_{1t} + b_{m2}Y_{mt} + ... + b_{mm}Y_{mt} + c_{m1}X_{1t} + c_{m2}X_{2t} + ... + c_{mk}X_{kt} = \varepsilon_{mt}
\end{aligned}
$$

يتم في المرحلة الأولى إجراء انحدار لكل متغير داخلي على جميع المتغيرات الخارجية الموجودة في النموذج الهيكلي، أي انحدار الصورة المختزلة:

$$
\begin{aligned}
& Y_{1t} = \alpha_{11}X_{1t} + \alpha_{12}X_{2t} + ... + \alpha_{1k}X_{kt} + u_{1t} \\
& Y_{2t} = \alpha_{21}X_{1t} + \alpha_{22}X_{2t} + ... + \alpha_{2k}X_{kt} + u_{2t} \\
& ..... \\
& Y_{mt} = \alpha_{m1}X_{1t} + \alpha_{m2}X_{2t} + ... + \alpha_{mk}X_{kt} + u_{mt}
\end{aligned}
$$

والذي يقود إلى قيمه مقدرة للمتغيرات الداخلية $\hat{Y}_{1t}, \hat{Y}_{2t}, ..., \hat{Y}_{mt}$

و في المرحلة الثانية، يتم استبدال المتغيرات الداخلية على يمين المعادلة الهيكلية بالقيم المقدرة ويعني ذلك استخدام القيم المقدرة (وتسمى متغير أداة Instrumental variable) بدلا عن القيم الحقيقية لتلك المتغيرات عند أجراء الانحدار و من تم يتم تقدير المعادلات التالية باستخدام طريقة المربعات الصغرى العادية:

$$
\begin{aligned}
& Y_{1t} = b_{12}\hat{Y}_{2t} + ... + b_{1m}\hat{Y}_{mt} + c_{11}X_{1t} + c_{12}X_{2t} + ... + c_{1k}X_{kt} + \varepsilon_{1t} \\
& Y_{2t} = b_{21}\hat{Y}_{1t} + b_{23}\hat{Y}_{3t} ... + b_{2m}\hat{Y}_{mt} + c_{21}X_{1t} + c_{22}X_{2t} + ... + c_{2k}X_{kt} + \varepsilon_{2t} \\
& ..... \\
& Y_{mt} = b_{m1}\hat{Y}_{1t} + b_{m2}\hat{Y}_{2t} + ... + c_{m1}X_{1t} + c_{m2}X_{2t} + ... + c_{mk}X_{kt} + \varepsilon_{mt}
\end{aligned}
$$

نلاحظ أن المتغير التابع مازال هو المتغير الداخلي الأصلي لكن التغيير تم في المتغيرات الداخلية الموجودة في الجانب الأيمن للمعادلة الهيكلية. من خواص المربعات الصغرى على مرحلتين:

- تتميز مقدرات 2SLS بأنها متسقة ولكن تظل متحيزة في العينات الصغيرة، فكلما كبر حجم العينة كلما كانت هذه المقدرات غير متحيزة
- يجب التأكد من المتغيرات الداخلة في النموذج المختزل، أي القيام باختبارات حسن التوفيق
- إذا كانت المتغيرات الخارجية مرتبطة فان النموذج لن يكون جيدا
- عند استخدام إحصائية Student لاختبار المعنوية الإحصائية للمعالم، مقدرات 2SLS أفضل بكثير من مقدرات OLS.

مثال 2:

لدينا البيانات الخاصة بالمتغيرات المذكورة في المثال 1 والمبينة في الجدول التالي، علما أن هذه المتغيرات مركزة:

الجدول (1)[1]: البيانات الإحصائية لـ Y_{1t}، Y_{2t}، X_t و $Y_{1,t-1}$

السنة	Y_{1t}	Y_{2t}	X_t	$Y_{1,t-1}$
1992	-30	-14	-26	20
1993	-4	10	5	-30
1994	-19	-19	-36	-4
1995	-6	-11	-6	-19
1996	-9	6	-13	-6
1997	11	-12	25	-9
1998	9	10	5	11
1999	19	11	32	9
2000	29	19	14	19

1- Bourbonnais (2003), p. 213

نذكر أن النموذج المبين في المثال 1 لا يحتوي على ثوابت لأن المعطيات مركزة، كما لا يمكن تطبيق طريقة المربعات الصغرى العادية على النموذج باعتبار أن عند تقدير المعادلة الهيكلية الثانية، نجد أن هناك تناقض، حيث أن Y_{1t} يظهر كمتغير مستقل بـ بالرغم من وجوده كمتغير تابع في المعادلة الأولى. فرضية استقلالية المتغير مع الخطأ غير محققة و بالتالي المقدرات التي سنحصل عليها ليست BLUE.

لتطبيق طريقة المربعات الصغرى غير المباشرة ILS، ينبغي أولا الانتقال من الشكل الهيكلي إلى الشكل المختزل، لدينا:

$$Y_{1t} = \frac{a}{1-ab}X_t + \frac{ac}{1-ab}Y_{1,t-1} + \frac{\varepsilon_{2t} + \varepsilon_{1t}}{1-ab}$$
$$Y_{2t} = \frac{ab}{1-ab}X_t + \frac{c}{1-ab}Y_{1,t-1} + \frac{\varepsilon_{2t} + b\varepsilon_{1t}}{1-ab}$$

أي:

$$Y_{1t} = \alpha_1 X_t + \beta_1 Y_{1,t-1} + \eta_{1t}$$
$$Y_{2t} = \alpha_2 X_t + \beta_2 Y_{1,t-1} + \eta_{2t}$$

نتائج التقدير بطريقة المربعات الصغرى العادية هي كالتالي:

$$\hat{Y}_{1t} = \underset{(4.93)}{0.717X_t} + \underset{(1.0)}{0.19Y_{1,t-1}}$$

$$R^2 = 0.78$$
$$n = 9$$

$$\hat{Y}_{2t} = \underset{(1.92)}{0.355X_t} + \underset{(1.0)}{0.139Y_{1,t-1}}$$

$$R^2 = 0.36$$
$$n = 9$$

حيث :(.) قيم ستيودنت

في هذه الحالة، النموذج على الشكل المختزل يكتب:

$$Y = -B^{-1}CX + B^{-1}\varepsilon = AX + \eta$$

مع: $A = -B^{-1}C$ هي مصفوفة معالم الشكل المختزل المقدرة

لدينا إذن: $BA = -C$

أي: $$\begin{pmatrix} 1 & -a \\ -b & 1 \end{pmatrix}\begin{pmatrix} 0.717 & 0.19 \\ 0.355 & 0.139 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix}$$

و الذي يعطينا:

$$0.717 - a \times 0.355 = a$$
$$0.190 - a \times 0.139 = 0$$
$$-b \times 0.190 + 0.355 = 0$$
$$-b \times 0.190 + 0.139 = c$$

يمكن من خلال المعادلتين الأخيرتين تقدير $\hat{b}$ و $\hat{c}$ حيث $\hat{b} = 0.495$ و $\hat{c} = 0.045$ لأن المعادلة الهيكلية الثانية معرفة تماما، بينما للمقدر $\hat{a}$ قيمتان أي أن المعامل a غير محدد لأن المعادلة الهيكلية الأولى زائدة تعريف. لهذا السبب، يجب تطبيـ ـق طريقـ ـة المربعـ ـات الصغرى على مرحلتين 2SLS باعتبار أن هذه الطريقة أكثر فعالية و أكثر دقة طالما هناك معادلة زائدة تعريف.

من أجل تقدير كل معالم النموذج، نستعين إذن بهذه الطريقة.

بالنسبة للمعادلة الأولى:

نقوم بانحدار المتغير الداخلي Y_{2t} على كل المتغيرات الخارجية في النموذج، ثم نقوم بحساب القيم المقدرة لـ . $\hat{Y}_{2t}$، حيث وجدنا في السابق:

$$\hat{Y}_{2t} = 0.355X_t + 0.139Y_{1,t-1}$$

نقوم بعدها بتعويض المتغير الداخلي Y_{2t} بمقدره $\hat{Y}_{2t}$ في المعادلة الأولى:

$$Y_{1t} = a(\hat{Y}_{2t} + X_t) + \varepsilon_{1t}$$

نتائج التقدير بطريقة المربعات الصغرى العادية على هذه الأخيرة هي كالتالي:

$$\hat{Y}_{1t} = 0.535(\hat{Y}_{2t} + X_t)$$

(5.22)

$$R^2 = 0.77$$
$$n = 9$$

أما بالنسبة للمعادلة الثانية:

نقوم بانحدار المتغير الداخلي Y_{1t} على كل المتغيرات الخارجية في النم وذج، ثم نق وم بحساب القيم المقدرة لـ . $\hat{Y}_{1t}$، حيث وجدنا في السابق:

$$\hat{Y}_{1t} = 0.717X_t + 0.19Y_{1,t-1}$$

ثم نعوض المتغير الداخلي Y_{1t} بمقدره $\hat{Y}_{1t}$ في المعادلة الثانية:

$$Y_{2t} = b\hat{Y}_{1t} + cY_{1,t-1} + \varepsilon_{2t}$$

نتائج التقدير بطريقة المربعات الصغرى العادية على هذه الأخيرة هي كالتالي:

$$\hat{Y}_{2t} = \underset{(1.921)}{0.495\hat{Y}_{1t}} + \underset{(0.181)}{0.045Y_{1,t-1}}$$
$$R^2 = 0.36$$
$$n = 9$$

مثال 3:

ليكن النموذج الاقتصادي الكلي التالي و المعدل من طرف Pindyck and Rubinfeld (1981)، الذي يعتبر من أحسن الأمثلة البيداغوجية للمشاكل التي يمكن أن نواجهه ا في إطار دراستنا للمعادلات الآنية:

$$C_t = a_0 + a_1Y_t + a_2C_{t-1} + \varepsilon_{1t}$$
$$I_t = b_0 + b_1I_{t-1} + b_2R_{t-1} + b_3Y_t + b_4i_{t-4} + \varepsilon_{2t}$$
$$i_t = c_0 + c_1Y_t + c_2R_t + c_3\nabla M_t + c_4is_{t-1} + \varepsilon_{3t}$$
$$Y_t = C_t + I_t + G_t$$
$$is_t = i_t + i_{t-1}$$
$$R_t = Y_t - Y_{t-1}$$

حيث:

C_t : الاستهلاك الإجمالي خلال الفترة t

Y_t : الناتج الوطني الإجمالي خلال الفترة t

I_t : الاستثمار خلال الفترة t

i_t : معدل الفائدة خلال الفترة t

G_t : الإنفاق الحكومي خلال الفترة t

∇ : الفروقات من الدرجة الأولى

M_t : عرض النقود خلال الفترة t

R_t : معدل النمو الاقتصادي خلال الفترة t

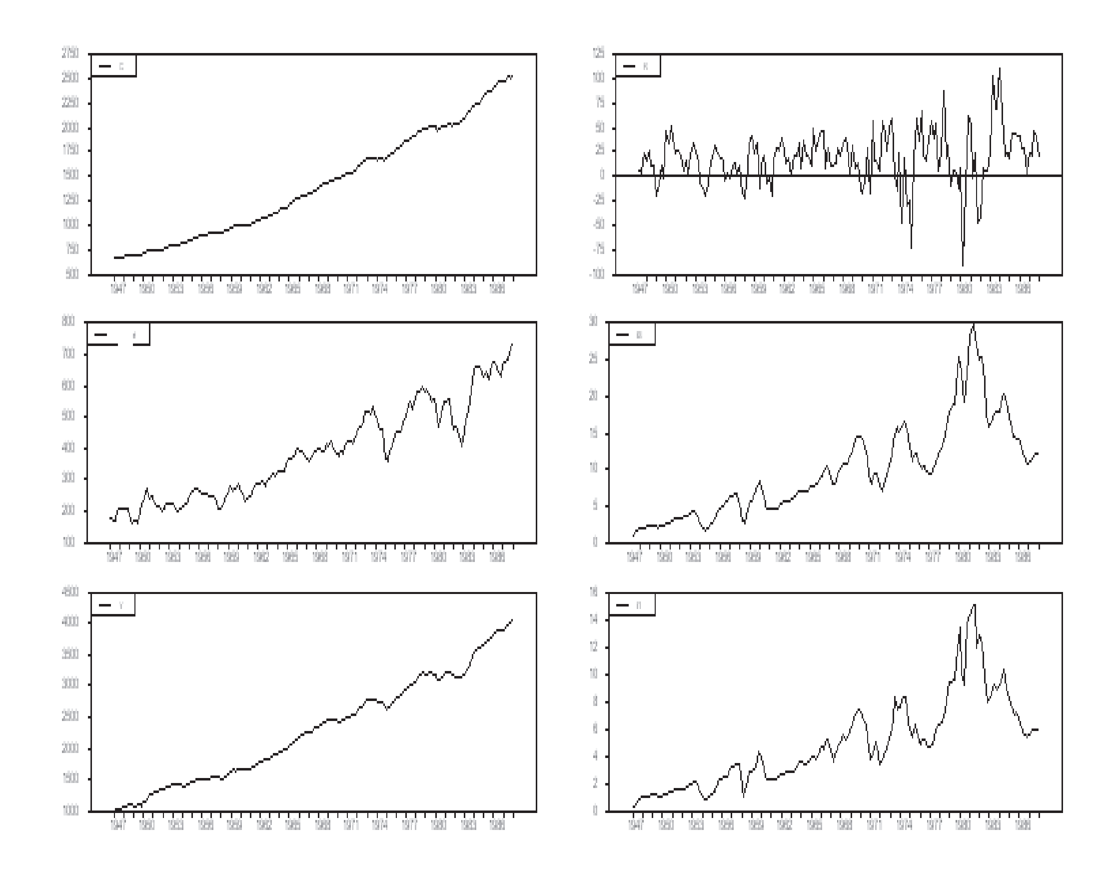
المطلوب:

1. تحديد المتغيرات الداخلية و الخارجية
2. تقدير المعادلات بطريقة المربعات الصغرى
3. تحديد شروط التمييز (أو التعريف) و الطريقة المناسبة للتقدير

الحل:

1. يحتوي النموذج على 6 متغيرات داخلية و هي C_t، I_t، i_t، Y_t، is_t و ∇Y_t، أي $m=6$ و 9 متغيرات خارجية C_{t-1}، G_t، ∇M_t، I_{t-1}، R_{t-1}، i_{t-4}، Y_{t-1}، i_{t-1} و 1 (الثوابت تعتبر معاملات متغير خارجي و هو الواحد)، أي $k=9$. نلاحظ أنه لا يوجد أي قيد خطي في النموذج، حيث $r=0$. تتم إذن هذه الدراسة خلال الفترة 1947 و 1988 على شكل معطيات فصلية، نمثل بيانيا كل المتغيرات الداخلية.

الشكل رقم (1): التمثيل البياني للمتغيرات الداخلية



2. نقوم بتقدير كل معادلة بطريقة المربعات الصغرى العادية، فنحصل على:

$$\hat{C}_t = -9.45 + 0.05Y_t + 0.92C_{t-1}$$
$$(-2.01)\ (3.20)\quad (36.73)$$
$$R^2 = 0.99,\ n = 144,\quad DW = 1.57$$

$$\hat{I}_t = -31.80 + 0.57I_{t-1} + 0.19R_{t-1} + 0.09Y_t - 5.65i_{t-4}$$
$$(-5.00)\ (12.18)\quad (3.70)\quad (8.90)\ (-5.83)$$
$$R^2 = 0.98\ ,\ n = 144\ ,\ DW = 1.85$$

$$\hat{i}_t = -0.55 + 0.0005Y_t + 0.013R_t - 0.085\nabla M_t + 0.42is_{t-1}$$
$$(-1.83)\ (2.24)\quad (4.31)\quad (-5.75)\quad (16.72)$$
$$R^2 = 0.93\ ,\ n = 144\ ,\ DW = 1.36$$

حيث أن القيم التي بين قوسين (.) هي إحصائيات ستيودنت.

من الملاحظ أن المقدرات المتحصل عليها بطريقة المربعات الصغرى العادية متحيـزة و بالتالي هذه الطريقة غير مناسبة لهذا النوع من النماذج و عليه يجب دراسة شروط التمييز و اختيار طريقة تقدير مناسبة.

3. بتطبيق شروط التمييز، نستنتج أن في المعادلة الهيكلية الأولى متغيرين داخليين $m^{'} = 2$ و متغيرين خارجين $k^{'} = 2$، هذا يعني أن:

$m - m^{'} + k - k^{'} + r = 6 - 2 + 9 - 1 + 0 = 12 > m - 1 = 6 - 1 = 5$ ومنه المعادلة الأولى زائدة تعريف، أما المعادلة الثانية، هناك أيضا متغيران داخليان و أربع متغيرات خارجية $m^{'} = 2$ و $k^{'} = 4$، أي:

$m - m^{'} + k - k^{'} + r = 6 - 2 + 9 - 4 + 0 = 9 > m - 1 = 6 - 1 = 5$، فهي إذن زائدة تعريف. بالنسبة للمعادلة الثالثة، هناك ثلاث متغيرات داخلية و ثلاث متغيرات خارجية، من الملاحظ أيضا أنها زائدة تعريف باعتبار أن

$m - m^{'} + k - k^{'} + r = 6 - 3 + 9 - 3 + 0 = 9 > m - 1 = 6 - 1 = 5$.

بما أن المعادلات زائدة تعريف، فلا يمكن تطبيق طرقة المربعات الصغرى غير المباشـرة ILS بل طريقة المربعات الصغرى المضاعفة أو طريقة المتغيرات الأداة. نقتـرح إذن هـذه الأخيرة باعتبار أنها سهلة و لا تحتاج إلى مجهود إضافي. النتائج مبينة باستعمال البرمجيـة RATS في الملحق.

6. مرحلة التنبؤ

بعد اختيار الطريقة المفضلة لتقدير كل معادلة من النموذج وفق للمعايير السابقة، يستخدم النموذج في التنبؤ إلا في حالة ما إذا صادفنا بعض المشاكل في القياس الاقتصادي مثل التعدد الخطي، الارتباط الذاتي بين الأخطاء أو عدم تجانس التباين، فإننا نقوم بتصحيح النموذج كما هو معروف قبل الخوض في عملية التنبؤ. قد يكون التنبؤ نقطيا punctual

أي توقع قيمة واحدة لكل متغير تابع في كل فترة مقبلة، و قد يكون بفترة أو مجال interval و يعني التنبؤ بمدى معين تقع داخله قيمة المتغير التابع باحتمال معين.

غير أن الإشكال المطروح هو أن التنبؤ بسلوك المتغيرات الداخلية خاضع للتنبؤ بسلوك المتغيرات الخارجية التي تفسرها في نفس الفترة و لن يتأتى ذلك إلا بدراسة التنبؤ عن طريق أسلوب السلاسل الزمنية. لقياس القدرة التنبؤية للنموذج الآني، نستعمل بعض المعايير الشائعة الاستخدام، نذكر منها متوسط مربعات الخطأ Mean Square Error (MSE) و متوسط الخطأ بالقيمة المطلقة Mean Absolute Error (MAE)،

$$MSE = H^{-1}\sum_{h=1}^{H}(Y_{i,n-H+h} - \hat{Y}_{i,n-H+h})^2$$

$$MAE = H^{-1}\sum_{h=1}^{H}\left|Y_{i,n-H+h} - \hat{Y}_{i,n-H+h}\right|$$

مع $i = 1,2,...,m$ و h هو أفق التنبؤ حيث $h = 1,2,...,H$. تعبر هذه المعايير عن مـ ـدى تأثر متوسط مربعات خطأ التنبؤ أو متوسط القيمة المطلقة لخطأ التنبؤ بالفرق بين القـ ـيم التنبؤية النظرية و تلك التقديرية. تجدر الإشارة إلى انه إذا كانت نسبة التحيز كبيرة، فـ ـإن هذا يدل على أن متوسط القيم المتوقعة تنحرف حقيقيا عن متوسط القيم الأصلية و هـ ـذا يعد خطأ يستدعي مراجعة بناء النموذج.

يمكن قياس كفاءة القدرة التنبؤية للنموذج باستعمال معامل تايل Theil و نرمز له بـ . U فيعبر عنه رياضيا كما يلي:

$$U = \frac{\sqrt{H^{-1}\sum_{h=1}^{H}(Y_{i,n-H+h} - \hat{Y}_{i,n-H+h})^2}}{\sqrt{H^{-1}\sum_{h=1}^{H}Y_{i,n-H+h}^2} + \sqrt{H^{-1}\sum_{h=1}^{H}\hat{Y}_{i,n-H+h}^2}}$$

و تنحصر قيمة U بين الصفر و الواحد الصحيح، فكلما اقتربت من الصفر أشار ذلـ ـك إلى كفاءة القدرة التنبؤية للنموذج.

مثال 4:

نأخذ معطيات المثال السابق، نذكر أن للنموذج ست متغيرات داخلية. لدراسة القدرة التنبؤية لهذا النموذج، نقوم بإعادة تقدير النموذج إلى غاية 1981 و من تم نحسب التنبؤ خلال الفترة الممتدة بين 1982 و 1985.

الجدول التالي يعطي نتائج اختبار القدرة التنبؤية:

الجدول (2): القدرة التنبؤية للنموذج

معامل U	MAE (%)	MSE (%)	h	المتغيرات الداخلية
0.49	10.23	12.50	1	C_t
0.26	12.01	12.90	2	
0.13	11.32	13.82	3	
0.18	10.75	13.33	4	
0.68	18.70	24.82	1	I_t
0.41	18.81	25.25	2	
0.29	18.42	26.13	3	
0.23	19.08	25.32	4	
1.14	0.74	1.14	1	i_t
0.67	0.78	0.93	2	
0.37	0.82	1.22	3	
0.54	0.65	1.18	4	
0.50	19.73	27.25	1	Y_t
0.26	19.74	27.69	2	
0.18	20.19	28.07	3	
0.14	19.62	28.94	4	
0.64	0.74	1.14	1	is_t
0.37	0.78	1.18	2	
0.28	0.82	1.22	3	
0.17	0.65	0.93	4	
0.84	19.73	27.25	1	R_t
0.48	19.74	28.94	2	
0.56	20.19	28.07	3	
0.61	19.62	27.69	4	

عند قراءتنا للنتائج، يتضح أن للنموذج قدرة تنبؤية كفؤة نوعا ما، حيث نلاحظ، باستثناء معدل الفائدة ومعدل النمو، أن قيم معامل U تتناقص نحو قيم تقترب من الصفر كلما زدنا في أفق التنبؤ، إلا أن التنبؤ الخاص بمعدل النمو غير جيد و هذا ما نلاحظه من خلال النتائج، فيتضح جليا أن معظم القيم تقترب من الواحد. تم الحصول على هذه النتائج وذلك بالاستعانة بالتعليمة الخاصة ببرمجية RATS المبينة في الملحق.

الملحق:

1. نقوم بإعطاء نتائج تقدير النموذج الآني المقترح في المثال الأول و الثاني باستعمال RATS وفق البرنامج التالي:

```
cal 1992 1 1
all 2000:1
data(org=obs,unit=input) / y1 y2 x ly
-30  -14   -26   20
-4    10    5   -30
-19  -19   -36   -4
-6   -11   -6   -19
-9     6   -13   -6
11   -12    25   -9
9     10    5    11
19    11    32    9
29    19    14   19
```

بالنسبة للمعادلة الأولى، نقوم بانحدار المتغير الداخلي Y_{2t} على كل المتغيرات الخارجية في النموذج، ثم نقوم بحساب القيم المقدرة لـ $\hat{Y}_{2t}$. و نسميها y2hat:

```
linreg y2 / res1
#x ly

Linear Regression - Estimation by Least Squares
Dependent Variable Y2
Usable Observations      9     Degrees of Freedom     7
Centered R**2     0.366987      R Bar **2   0.276556
Uncentered R**2   0.366987      T x R**2       3.303
Mean of Dependent Variable       0.000000000
Std Error of Dependent Variable 13.874436926
Standard Error of Estimate      11.800970018
Sum of Squared Residuals        974.84025353
Durbin-Watson Statistic             3.010084

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  X                        0.3553575307 0.1849568379      1.92130  0.09614568
2.  LY                       0.1396826998 0.2431053326      0.57458  0.58356267


set y2hat / = y2(t)-res1(t)
```

نقوم بعدها بتعويض المتغير الداخلي Y_{2t} بمقـدره $\hat{Y}_{2t}$ في المعادلـة الأولى و بمـا أن المتغيرين $\hat{Y}_{2t}$ و X_t لهما نفس المعامل، فيمكن استبدال مجموعهما بمتغير آخر Z_t و عليه نقدر المعادلة الأولى:

```
set z / = y2hat(t)+x(t)


linreg y1 / resids1
#z


Linear Regression - Estimation by Least Squares
Dependent Variable Y1
Usable Observations      9     Degrees of Freedom     8
Centered R**2      0.773287      R Bar **2   0.773287
Uncentered R**2    0.773287      T x R**2       6.960
Mean of Dependent Variable       0.000000000
Std Error of Dependent Variable 18.701604209
Standard Error of Estimate       8.904647886
Sum of Squared Residuals       634.34203186
Durbin-Watson Statistic            1.638195

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Z                      0.5355007515 0.1025138470      5.22369  0.00079899
```

نفس الشيء بالنسبة للمعادلة الثانية، نقوم بانحدار المتغير الـداخلي Y_{1t} علـى كـل المتغيرات الخارجية في النموذج، ثم نقوم بحساب القيم المقدرة لـ $\hat{Y}_{1t}$. و نسميها y1hat:

```
linreg y1 / res2
#x ly

Linear Regression - Estimation by Least Squares
Dependent Variable Y1
Usable Observations      9     Degrees of Freedom     7
Centered R**2     0.784723      R Bar **2   0.753970
Uncentered R**2   0.784723      T x R**2       7.063
Mean of Dependent Variable       0.000000000
Std Error of Dependent Variable 18.701604209
Standard Error of Estimate       9.276267758
Sum of Squared Residuals       602.34400467
Durbin-Watson Statistic            1.557471

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  X                       0.7174487379 0.1453871292      4.93475  0.00168516
2.  LY                      0.1909747828 0.1910953214      0.99937  0.35090145


set y1hat / = y1(t)-res2(t)
```

نعوض المتغير الداخلي Y_{1t} بمقدره $\hat{Y}_{1t}$ في المعادلة الأولى، ثم نقدر:

```
linreg y2 / resids2
#y1hat ly

Linear Regression - Estimation by Least Squares
Dependent Variable Y2
Usable Observations      9     Degrees of Freedom     7
Centered R**2     0.366987      R Bar **2   0.276556
Uncentered R**2   0.366987      T x R**2       3.303
Mean of Dependent Variable       0.000000000
Std Error of Dependent Variable 13.874436926
Standard Error of Estimate      11.800970018
Sum of Squared Residuals       974.84025353
Durbin-Watson Statistic            3.010084

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Y1HAT                   0.4953072072 0.2577979835      1.92130  0.09614568
2.  LY                      0.0450915135 0.2488180452      0.18122  0.86132830
```

نلاحظ أن هذه الطريقة طويلة و تأخذ مجهودا إضافيا و لهذا السبب، تسمح البرمجيات باستعمال هذه الطريقة وذلك بتعليمة واحدة، فيتعلق الأمر هنا بطريقة المتغيرات الأداة.

التعليمة هي كالتالي:

- المعادلة الأولى:

```
INSTRUMENTS X LY
LINREG(INST) Y1
#Y2 X
```

- المعادلة الثانية:

```
INSTRUMENTS X LY
LINREG(INST) Y2
#Y1 LY
```

2. نعطي نتائج تقدير النموذج الآني المقترح في المثال الثالث:

```
open data prsmall.xls
cal 1947 1 4
all 1988:1
data(format=xls,org=obs)
```

نمثل بيانيا كل مشاهدات المتغيرات الداخلية و ذلك بالاستعانة بالبرنامج التالي:

```
SPGRAPH(HFIELDS=3,VFIELDS=2)
 DOFOR SERIES = CONS INVEST RATE R IS Y
  GRAPH(KEY=LOLEFT) 1
  # SERIES
 END DO
SPGRAPH(DONE)
```

نقدر المعادلة الهيكلية الأولى بطريقة المربعات الصغرى العادية:

```
LINREG CONS 1950:1 1985:4
#CONSTANT GNP CONS{1}

Linear Regression - Estimation by Least Squares
Dependent Variable CONS
Quarterly Data From 1950:01 To 1985:04
Usable Observations    144     Degrees of Freedom   141
Centered R**2     0.999475      R Bar **2   0.999468
Uncentered R**2   0.999945      T x R**2     143.992
Mean of Dependent Variable      1411.1625000
Std Error of Dependent Variable  486.7321052
Standard Error of Estimate        11.2309913
Sum of Squared Residuals       17785.058456
Regression F(2,141)             134221.4081
Significance Level of F           0.00000000
Durbin-Watson Statistic             1.573354

   Variable                    Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant               -9.454363868  4.703227512     -2.01019  0.04631886
2.  GNP                     0.054095564  0.016865400      3.20749  0.00165727
3.  CONS{1}                 0.925963960  0.025206774     36.73473  0.00000000
```

نفس الشيء بالنسبة للمعادلة الثانية:

```
LINREG INVEST 1950:1 1985:4
# CONSTANT INVEST{1} R{1} GNP RATE{4}

Linear Regression - Estimation by Least Squares
Dependent Variable INVEST
Quarterly Data From 1950:01 To 1985:04
Usable Observations    144     Degrees of Freedom   139
Centered R**2     0.984317      R Bar **2   0.983866
Uncentered R**2   0.998372      T x R**2     143.766
Mean of Dependent Variable      383.83541667
Std Error of Dependent Variable 131.09401018
Standard Error of Estimate       16.65147819
Sum of Squared Residuals       38540.769925
Regression F(4,139)               2181.0786
Significance Level of F           0.00000000
Durbin-Watson Statistic             1.856869

   Variable                    Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant               -31.80043760   6.35198259     -5.00638  0.00000165
2.  INVEST{1}                0.57520823   0.04719432     12.18808  0.00000000
3.  R{1}                     0.19929317   0.05384506      3.70123  0.00030837
4.  GNP                      0.09594195   0.01077263      8.90608  0.00000000
5.  RATE{4}                 -5.65550354   0.96966562     -5.83243  0.00000004
```

و المعادلة الثالثة، كما يلي:

```
LINREG RATE 1950:1 1985:4
# CONSTANT GNP R MDIFF IS{1}

Linear Regression - Estimation by Least Squares
Dependent Variable RATE
Quarterly Data From 1950:01 To 1985:04
Usable Observations    144     Degrees of Freedom   139
Centered R**2     0.933684      R Bar **2   0.931776
Uncentered R**2   0.981067      T x R**2     141.274
Mean of Dependent Variable      5.1534027778
Std Error of Dependent Variable 3.2689143326
Standard Error of Estimate      0.8538311114
Sum of Squared Residuals        101.33483178
Regression F(4,139)                 489.2595
Significance Level of F           0.00000000
Durbin-Watson Statistic             1.367418

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                -0.556221382  0.303526543     -1.83253  0.06901262
2.  GNP                      0.000512001  0.000228540      2.24031  0.02665742
3.  R                        0.013453448  0.003117956      4.31483  0.00003012
4.  MDIFF                   -0.085249291  0.014822048     -5.75152  0.00000005
5.  IS{1}                    0.425933095  0.025465197     16.72609  0.00000000
```

كما لاحظنا أن هذه الطريقة تعطي تقديرات متحيزة وعليه نستعمل طريقة المتغيرات الأداة، لتكن نتائج تقدير المعادلات الثلاثة على الترتيب و المبينة في مخرجات البرمجية RATS:

```
INSTRUMENTS CONSTANT CONS{1} R{1} GNP{1} INVEST{1} G MDIFF RATE[1} RATE{4}
LINREG(INST,FRML=CONSEQ) CONS 1950:1 1985:4
# CONSTANT GNP CONS{1}

Linear Regression - Estimation by Instrumental Variables
Dependent Variable CONS
Quarterly Data From 1950:01 To 1985:04
Usable Observations    144     Degrees of Freedom   141
Centered R**2     0.999464      R Bar **2   0.999457
Uncentered R**2   0.999943      T x R**2     143.992
Mean of Dependent Variable      1411.1625000
Std Error of Dependent Variable  486.7321052
Standard Error of Estimate        11.3436190
Sum of Squared Residuals        18143.554646
J-Specification(6)                 57.925273
Significance Level of J           0.00000000
Durbin-Watson Statistic             1.630761

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                 -3.179856589  4.955717813     -0.64165  0.52213954
2.  GNP                       0.025662686  0.018196110      1.41034  0.16064182
3.  CONS{1}                   0.968332903  0.027185647     35.61927  0.00000000


LINREG(INST,FRML=INVESTEQ) INVEST 1950:1 1985:4
# CONSTANT INVEST{1} R{1} GNP RATE{4}

Linear Regression - Estimation by Instrumental Variables
Dependent Variable INVEST
Quarterly Data From 1950:01 To 1985:04
Usable Observations    144     Degrees of Freedom   139
Centered R**2     0.984092      R Bar **2   0.983634
Uncentered R**2   0.998349      T x R**2     143.762
Mean of Dependent Variable      383.83541667
Std Error of Dependent Variable 131.09401018
Standard Error of Estimate       16.77065169
Sum of Squared Residuals        39094.411355
J-Specification(4)                 25.495441
Significance Level of J           0.00003999
Durbin-Watson Statistic             2.004856

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                 -25.50262772   6.43431018     -3.96354  0.00011761
2.  INVEST{1}                  0.63725350   0.04801265     13.27262  0.00000000
3.  R{1}                       0.21449446   0.05425583      3.95339  0.00012218
4.  GNP                        0.08071954   0.01097636      7.35394  0.00000000
5.  RATE{4}                   -4.66692375   0.98255518     -4.74978  0.00000501
```

```
LINREG(INST,FRML=RATEEQ) RATE 1950:1 1985:4
# CONSTANT GNP R MDIFF IS{1}

Linear Regression - Estimation by Instrumental Variables
Dependent Variable RATE
Quarterly Data From 1950:01 To 1985:04
Usable Observations    144     Degrees of Freedom   139
Centered R**2     0.923342      R Bar **2   0.921136
Uncentered R**2   0.978115      T x R**2     140.849
Mean of Dependent Variable      5.1534027778
Std Error of Dependent Variable 3.2689143326
Standard Error of Estimate      0.9179997089
Sum of Squared Residuals        117.13856171
J-Specification(4)                 56.214944
Significance Level of J           0.00000000
Durbin-Watson Statistic             1.450961

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                 0.273120392  0.342410955      0.79764  0.42644042
2.  GNP                     -0.000434281  0.000276810     -1.56888  0.11895023
3.  R                        0.022534344  0.005184064      4.34685  0.00002650
4.  MDIFF                   -0.079425302  0.017715078     -4.48349  0.00001523
5.  IS{1}                    0.541787141  0.031239423     17.34306  0.00000000
```

لدراسة كفاءة التنبؤ بقيم المتغيرات الداخلية في النموذج الآني، نستعمل أيضا البرمجيـة RATS لهذا الغرض و البرنامج هو كالتالي:

```
THEIL(SETUP,MODEL=PRSMALL) 6 4 1985:4
DO TIME=192:1,1985:4
  THEIL TIME
END DO TIME
THEIL(DUMP,WINDOW='Forecast Performance')
```

حيث 6 يعبر عن عدد المعادلات في النموذج و 4 عدد التنبؤات و هنا $h = 4$.

الفصل السادس

أدوات تحليل السلاسل الزمنية

الفصل السادس

أدوات تحليل السلاسل الزمنية

تناولنا في الفصول السابقة شكلا من أشكال النمذجة القياسية، ويتمثل في نماذج الانحدار، التي تعتمد في تفسيرها للظاهرة على عدد من المتغيرات المستقلة التي تؤثر فيها، وكيفية صياغة هذه النماذج. أما في هذا الفصل سنتناول شكلا آخر يتمثل في نماذج السلاسل الزمنية الخطية، التي تعتمد في تفسيرها للظاهرة في اللحظة الحالية على المتوسطات المرجحة للملاحظات الماضية والأخطاء العشوائية. ويُشترط في هذا الشكل أن تكون السلسلة مستقرة. لتحقق هذه الصفة من عدمه يوجد عدة اختبارات إحصائية مخصصة لذلك. قبل التطرق إلى منهجية Box-Jenkins للتنبؤ، سنقوم بالتمييز بين عدة أنواع من النماذج الخطية للسلاسل الزمنية AR, MA, ARMA, ARIMA, SARIMA.

1. تعريف السلسلة الزمنية و مركباتها:

السلسلة الزمنية هي مجموعة من القيم لمؤشر إحصائي معين مرتبة حسب تسلسل زمني، بحيث كل فترة زمنية يقابلها قيمة عددية للمؤشر تسمى مستوى السلسلة. وبمعنى آخر هي مجموعة من المعطيات ممثلة عبر الزمن المرتب ترتيبا تصاعديا[1].

عند بناء السلسلة الزمنية، وقبل استخدامها في التحليل أو التنبؤ، لا بد من التأكد أن مستوياتها قابلة للمقارنة فيما بينها، وهو شرط أساسي لصحة أي تحليل وأي تقدير وأي توقع. يشترط أن تكون جميع مستويات السلسلة خاصة بمكان معين، سواء أ كان إقليما أو ولاية أو مؤسسة[2] وأن تكون وحدة القياس لجميع مستويات السلسلة الزمنية موحدة. تجدر الإشارة إلى أن السلاسل الزمنية عادة ما لا تُعطَى جاهزة وقابلة للتحليل مباشرة،

1- David and Michaud, 1989, p.22.

2- عبد العزيز شرابي، 2000، ص 20.

حيث يتطلب الأمر في أغلب الأحيان إجراء بعض التعديلات لجعل المستويات قابلة للمقارنة.

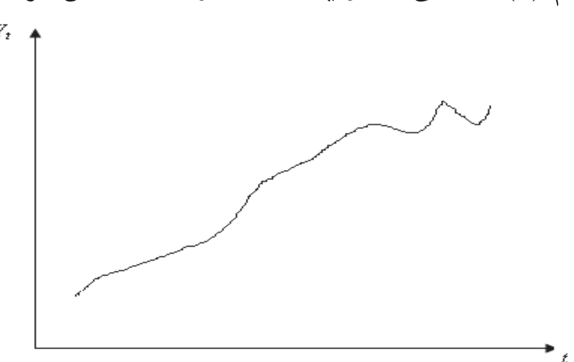
تتكون السلسلة الزمنية من مجموعة من المركبات التي تساعدنا على معرفة سلوك السلسلة وتحديد مقدار تغيراتها وإدراك طبيعتها واتجاهها حتى يصبح بالإمكان القيام بالتقديرات اللازمة والتنبؤات الضرورية، وهذه العناصر هي:

1.1. الاتجاه العام Trend:

هو النمو الطبيعي للظاهرة، حيث يعبر عن تطور متغير ما عبر الزمن، سواء أ كان هذا التطور بميل موجب أو سالب، إلا أن هذا التطور لا يُلاحظ في الفترات القصيرة، بينما يكون واضحا في الفترات الطويلة ويرمز له بالرمز T_t. تكون مشاهدات السلسلة الزمنية تابعة للزمن الذي يحدد خاصيتها أو سمتها الرئيسية، وهذه العلاقة الزمنية قد تأخذ أشكالا مختلفة.

والشكل البياني التالي يوضح حالة وجود مركبة اتجاه عام في السلسلة الزمنية Y_t:

الشكل رقم (1): منحنى معياري لسلسلة زمنية تتضمن مركبة اتجاه عام



2.1. التغيرات الموسمية Seasonal variations:

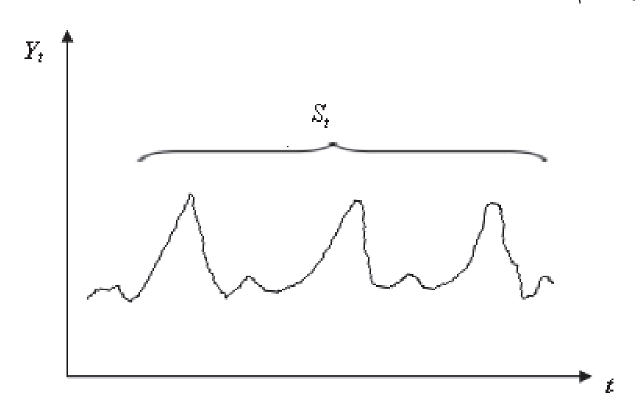
هي التغيرات التي تحدث بانتظام في وحدات زمنية متعاقبة والتي تنجم من تأثير عوامل خارجية، أو هي تقلبات قصيرة المدى تتكرر على نفس الوتيرة كل سنة[1]، ويرمز لها بـ . .

1- Grais, 1978, p. 326.

S_t. وكمثال لهذه التغيرات العطل والإجازات، الإقبال على نوع من الألبسة في فصل ما، استهلاك المكيف في فصل الصيف...الخ.

والشكل التالي يوضح حالة وجود مركبة موسمية في السلسلة الزمنية Y_t:

الشكل رقم (2): منحنى معياري لسلسلة زمنية تتضمن مركبة موسمية

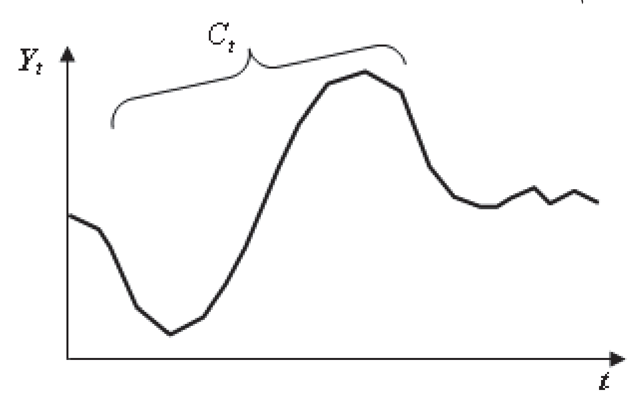


3.1. التغيرات الدورية Cyclical Variations:

تنعكس هذه المركبة في السلاسل الزمنية طويلة المدى، والتي تبرز انتقال أثر الأحـ ـوال الاقتصادية مثلا، وهي تغيرات تشبه التغيرات الموسمية إلا أنها تتم في فترات أطول نسبيا من الفترات الموسمية، وبالمقارنة بالتغيرات الموسمية فإن طول الفترة الزمنية غير معلـ ـوم وإنمـ ـا يتراوح عادة بين ثلاث سنوات إلى عشر سنوات، وبالتالي يصعب التعرف على التقلبات الدورية ومقاديرها لأنها تختلف اختلافا كبيرا من دورة لأخرى سواء من حيث طول الفترة الزمنية للدورة أو اتساع تقلباتها ومداها، ونرمز لها بالرمز C_t.

والشكل البياني التالي يوضح حالة وجود مركبة الدورات في السلسلة الزمنية Y_t:

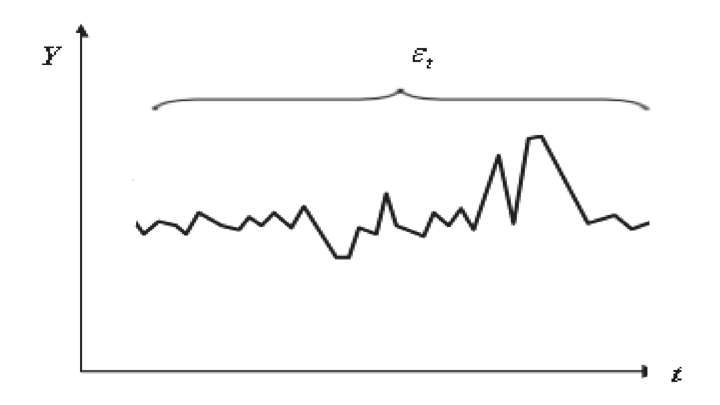
الشكل رقم (3): منحنى معياري لسلسلة زمنية تتضمن مركبة دورية



4.1. التغيرات العشوائية Random or Stochastic variations:

وهي تعبر عن تلك التذبذبات غير المنتظمة، و بمعنى أخر هي تلك التغيرات الشاذة التي تنجم عن ظروف طارئة لا يمكن التنبؤ بوقوعها أو تحديد نطاق تأثيرها، حيث تنشأ عن أسباب عارضة لم تكن في الحسبان مثل الزلازل، إضراب العمال...الخ، ويرمز لها بـ: ε_t. والشكل التالي يوضح حالة وجود مركبة عشوائية في السلسة الزمنية Y_t:

الشكل رقم (4): منحنى يبين التغيرات العشوائية في السلسلة الزمنية



لكي نستطيع إجراء تحليل السلاسل الزمنية إلى مركباتها يجب أن يكون لدينا نموذج لها، وهذا يعني أن نحدد العلاقة بين مكونات السلسلة الزمنية، وهناك نموذجان شائعا الاستخدام:

أ- نموذج الجمع: $Y_t = T_t + S_t + C_t + \varepsilon_t$

ب- نموذج الجداء: $Y_t = T_t \times S_t \times C_t \times \varepsilon_t$

ويمكن معرفة طبيعة النموذج انطلاقا من حساب المتوسط الحسابي والانحراف المعياري، فإذا كان هذان الأخيران ثابتين عبر وحدة الزمن (مستقلين) فإن السلسلة تشكل نموذجا تجميعيا، وفي حالة العكس نقول عن السلسلة أنها تشكل نموذجا جدائيا[1]، وعند إجراء تعديلات على النموذج الجدائي نحصل على نموذج تجميعي ويتم تحليل السلاسل الزمنية لعزل المؤثرات المنتظمة وغير المنتظمة، ومعرفة مدى تأثير كل منها على قيمة الظاهرة المشاهدة وبذلك يكون القصد من التحليل رد القيمة الكلية للظاهرة إلى عناصرها المكونة لها[2].

يمكن كشف وجود مركبات السلاسل الزمنية عن طريق تحليل المعلومات بيانيا، فيتمثل الاتجاه العام في تلك المركبة التي تدفع بمنحنى تطور السلسلة عبر الزمن إلى الأعلى (ميل موجب)، أو إلى الأسفل (ميل سالب)، بينما تنعكس المركبة الدورية في الشكل البياني على هيئة قِمم أو انخفاضات بشكل منتظم يسمح لنا بتحديد فترة حدوث هذه الظاهرة. وأما المتغيرة العشوائية تتمثل في التذبذب الحاصل على مستوى السلسلة، أما المتغيرة الموسمية تتضح من خلال الانتظام الموجود في تسجيل قيمة على الفصل الأخير لكل سنة، أو انخفاض في كل بداية سنة جديدة مثلا. وإلى جانب التحليل البياني يوجد عدة اختبارات إحصائية مخصصة لكشف هذه المركبات منها اختبار دانيال لكشف مركبة الاتجاه العام حيث يعتبر هذا الأخير من أهم المركبات التي تتكون منها السلسلة الزمنية وذلك لأنها تستخدم في عمليات التنبؤ بقيم الظاهرة للفترات الزمنية المستقبلية، ويمكن

1- Bourbonnais and Terraza (1998), p. 15.

2- علي لزعر، 2000، ص 141.

تقدير هذه المركبة بعدة طرق منها التمهيد باليد، طريقة الأوساط المتحركة للتخلص من الذبذبات الموسمية، حتى يظهر بوضوح الاتجاه العام للظاهرة محل الدراسة، كما يمكن استخدام طريقة المربعات الصغرى[1]. هناك أيضا اختبار كريسكال والس Kruskall-Wallis لكشف المركبة الموسمية و لكن هذه الاختبارات تعتبر غير فعالة بالمقارنة مع اختبار الجذر الوحدوي الذي يمكننا من معرفة وجود اتجاه عام أو مركبة موسمية.

2. السلاسل الزمنية المستقرة وغير المستقرة:

قبل الشروع في دراسة تقلبات أي ظاهرة اقتصادية لا بد من التأكد أولا من وجود اتجاه في السلسلة الزمنية، وحسب طبيعة نمو السلسلة يمكننا أن نميز بين سلاسل زمنية مستقرة Stationary Time Series، وسلاسل زمنية غير مستقرة Non stationary Time Series أي ذات اتجاه.

كون السلسلة تحمل هذه الخاصية أو تلك لها علاقة مباشرة باختيار تقنية التوقع المناسبة، وهناك حتى من يُصنِّف تقنيات التوقع على هذا الأساس (مستقرة أو غير مستقرة). إن السلسلة الزمنية المستقرة هي تلك التي تتغير مستوياتها مع الزمن دون أن يتغير المتوسط فيها، وذلك خلال فترة زمنية طويلة نسبيا، أي أن السلسلة لا يوجد فيها اتجاه لا نحو الزيادة ولا نحو النقصان، أما السلسلة الزمنية غير المستقرة فإن المستوى المتوسط فيها يتغير باستمرار سواء نحو الزيادة أو النقصان، وهذا تمثيل بياني لسلسلة زمنية غير مستقلة.

نقول على سلسلة زمنية ما مستقرة بمعنى ضعيف Wide sense stationarity، إذا كانت توقعها، تباينها، وتبايناتها المشتركة ثابتة عبر الزمن أي[2]:

1. تذبذبت حول متوسط حسابي ثابت عبر الزمن: $E(Y_t) = E(Y_{t+k}) = \mu$
2. ثبات التباين عبر الزمن:

1- عبد الرحمن بن محمد سليمان أبو عمه، أنور أحمد محمد عبد الله، محمود محمد إبراهيم هنيدي، 1995، ص 197.
2- تومي صالح، 1999، ج(2)، ص 173.

$$\operatorname{var}(Y_t) = E[Y_t - E(Y_t)]^2 = \operatorname{var}(Y_{t+k}) = E[Y_{t+k} - E(Y_{t+k})]^2 = \gamma(0) = \sigma^2 < \infty \ , \ \forall t$$

3. أن يكون التباين المشترك بين أي قيمتين لنفس المتغير معتمدا على الفجوة الزمنية بين القيمتين، وليس على القيمة الفعلية للزمن الذي يحسب عند التغاير، أي على الفرق بين فترتين زمنيتين.

$$\operatorname{cov}(Y_t, Y_{t+k}) = E[(Y_t - \mu)(Y_{t+k} - \mu)] = \operatorname{cov}(Y_{t+k}, Y_{t+k+s}) = \gamma(k)$$

قد يصعب أحيانا تحديد طبيعة السلسلة الزمنية (مستقرة أو غير مستقرة) سواء بالملاحظة البسيطة أو حتى بالرسم البياني، هنا نلجأ إلى استخدام مقاييس إحصائية لاختبار وجود أو عدم وجود الاتجاه العام في السلسلة، أبسط هذه المقاييس وأكثرها استعمالا هي القيام بتقسيم السلسلة الزمنية إلى قسمين متساويين ثم حساب المتوسط الحسابي لكل قسم، فإذا كان المتوسطان الحسابيان متساويين أو قريبين من بعضهما، نقول أنه لا يوجد اتجاه في السلسلة الزمنية وبالتالي فهي مستقرة، أما إذا كان هناك عدم تساوي ملحوظ فإننا نستنتج أن هناك اتجاه، أي أن السلسلة الزمنية غير مستقرة، ويمكن التأكد أكثر وذلك باختبار معنوية هذا الاختلاف، (أي التأكد من أن الاختلاف بين المتوسطين معنوي و لم يكن نتيجة الصدفة). هناك أدوات مهمة في تحليل السلاسل الزمنية واختبار استقراريتها، هي دالة الارتباط الذاتي النظرية التي تختلف باختلاف النماذج وتساعد على تمثيل السلاسل الزمنية ميدانيا و أيضا اختبار الجذر الوحدوي الذي يعتبر الأداة الأكثر فعالية.

ولنأخذ مثالا يتمثل في نموذج السير العشوائي (Random Walk) :

$$Y_t = Y_{t-1} + \varepsilon_t$$
$$\varepsilon_t \sim IID\left(0, \sigma_\varepsilon^2\right) \qquad t = 1.....T$$

نستطيع التأكد من حالة عدم الاستقرار عن طريق إيجاد التباين والتباينات المشتركة للسلسلة واختبار ثباتها، وبأخذ التوقع نجد: $E(Y_t)=E(Y_{t-1})+E(\varepsilon_t)=E(Y_{t-1})$ ، ويعني ذلك أن هذه السيرورة مستقرة بالنسبة للمتوسط، أما بالنسبة للتباين فنجد:

$$\operatorname{var}(Y_t)=\operatorname{var}(Y_{t-1})+\operatorname{var}(\varepsilon_t)+2\operatorname{cov}(Y_{t-1},\varepsilon_t)$$

ومع ε_t مستقلة ومتماثلة التوزيع، يكون الحد الأخير للمعادلة أعلاه معدوما، فنجد في الأخير أن: $\operatorname{var}(Y_t)=\operatorname{var}(Y_{t-1})+\sigma_\varepsilon^2=t\sigma_\varepsilon^2$ أي أن $\operatorname{var}(Y_t)\neq\operatorname{var}(Y_{t-1})$ و التباين مرتبط بالزمن، ومنه فإن السير العشوائي غير مستقر بالنسبة للتباين، وهذا كاف لكي تكون السلسلة غير مستقرة، أما إذا أضفنا حدا ثابتا للسير العشوائي فالسيرورة غير مستقرة بالنسبة للمتوسط كذلك، وبالرغم من ذلك فإن الفروقات الأولى لـ Y_t. تعطي سيرورة مستقرة، ما دام: $\nabla Y_t=Y_t-Y_{t-1}=\varepsilon_t$.

حيث أن ε_t ذو تشويش أبيض White Noise، وهذا يبين أنه من الممكن في بعض الحالات تخصيص نموذج نظري غير مستقر، ثم نعمم سيرورة مستقرة بواسطة الفروقات، ومع هذا فهناك عدة حالات يكون فيها تطبيق الفروقات على السلسلة المستقرة لا يعطي نموذجا مستقرا، وتسمى هذه الحالات بالسلاسل غير المستقرة وغير المتجانسة، أما بالنسبة لنموذج السير العشوائي أعلاه فهو سيرورة غير مستقرة أصلا و متجانسة من الدرجة الأولى.

إن تخصيص نموذج غير مستقر تبعا للفروقات يمكن أن يؤدي بنا إلى عدة مشاكل مثل الحصول على سيرورة لا يمكن تفريقها من أجل الوصول إلى الاستقرار ولهذا نفهم لماذا يستخدم منمذجو السلاسل الزمنية عدة طرق، بواسطة حساب الفروقات للسلسلة الميدانية، من أجل الحصول على سلسلة محولة تظهر بأنها مستقرة، ثم ينظرون إلى النموذج الذي يمكن أن يمثل السلسلة المستقرة، ومنه نواجه مشكل كيفية التعرف على السلسلة الميدانية هل هي مستقرة أم لا ؟.

إن أول شيء نقوم به هو النظر إلى الرسم البياني للسلسلة، فإذا لاحظنا بوضوح تصاعد (أو تنازل) في الاتجاه العام للسلسلة تكون متوسطات مختلف العينات الجزئية للسلسلة مختلفة، وهذا يعني عدم إمكانية تعميم الملاحظات على سيرورة مستقرة، والتي تستلزم نفس قيمة المتوسط $E(Y_t)$ بالنسبة لكل فترة t، أي أن $E(Y_t)$ غير ثابت بالنسبة للزمن، وإذا فشلنا في تحديد استقرار السلسلة الميدانية من الرسم البياني، يمكن أن ننظر إلى دالة الارتباط الذاتي للعينة أو اختبار الجذر الوحدوي.

1.2. اختبار معنوية معاملات الارتباط الذاتي:

توضح دالة الارتباط الذاتي لسلسلة زمنية الارتباط الموجود بين المشاهدات لفترات مختلفة وهي ذات أهمية بالغة في إبراز بعض الخصائص الهامة للسلسلة الزمنية، ومن الناحية العملية نقوم بتقدير دالة الارتباط الذاتي للمجتمع بواسطة دالة الارتباط الذاتي للعينة، حيث تتمثل دالة الارتباط الذاتي عند الفجوة k كما يلي[1]:

$$\hat{\rho}(k)_k = \frac{\sum_{t=k+1}^{T}\left(Y_t - \bar{Y}\right)\left(Y_{t-k}\bar{Y}\right)}{\sum_{t=1}^{T}\left(Y_t - \bar{Y}\right)^2} \qquad ,t = 1,2,3,\ldots,T$$

ويمكن حساب الصيغة من بيانات عينة على النحو التالي:

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)}$$

حيث:

$$\hat{\gamma}(k) = \frac{\sum\left(Y_t - \bar{Y}\right)\left(Y_{t+k} - \bar{Y}\right)}{T-k}$$

و

$$\hat{\gamma}(0) = \frac{\sum\left(Y_t - \bar{Y}\right)^2}{T}$$

حيث T تمثل حجم العينة و k طول الفجوة الزمنية، وتتراوح قيمة معامل الارتباط الذاتي $\rho(k)$ بين -1 و +1.

1- Michel, 1994, p. 101.

نقول إذن عن سلسلة أنها مستقرة إذا كان معامل الارتباط الذاتي يساوي الصفر أو قريب منه لأي فجوة أكبر من الصفر، أي أنه في هذه الحالة يجب أن تنخفض الارتباطات الذاتية بسرعة كلما ارتفع k ، أما إذا كانت السلسلة غير مستقرة، فإن الخطوة القادمة هي محاولة تفريقها، لهدف الحصول على سلسلة محولة ومستقرة، وباستعمال W_t كأنه سلسلة مفرقة، يكون لدينا: $W_t = \nabla Y_t = Y_t - Y_{t-1} \quad ,t = 2,3......,T$

بعد استعمال الفروقات للسلسلة، يمكن النظر إلى كل من الرسم البياني للسلسة المفرقة ودالة الارتباط الذاتي، لهدف التأكد من عدم وجود مشكل عدم الاستقرار. إذا بقيت W_t غير مستقرة نواصل حساب الفروقات على الشكل:

$$W_t = \nabla^2 Y_t \quad , \quad t = 3,4......,T$$

ومنه يمكن أن نطبق عامل الفروقات d Differentiation coefficient مرة واحدة على السلسلة المشتقة:

$$W_t = \Delta^d Y_t \quad ,t = d+1, d+2,......,T$$

لكن عند تحليل دوال الارتباط الذاتي لسلسلة زمنية فإن السؤال الذي يطرح هو تحديد $\rho(k)$ التي تكون معنويا تختلف عن الصفر، بمعنى اختبار الفرضيتين:

$$H_0 : \rho(k) = 0$$
$$H_1 : \rho(k) \neq 0$$

إذ نستطيع استعمال معامل الارتباط الذي يرتكز على إحصائية ستيودنت، هذا من جهة، ومن جهة أخرى برهن Qunennouille (1949) على أنه من أجل $T \geq 30$ فإن المعامل $\rho(k)$ ينتهي تقاربيا إلى القانون الطبيعي ذي متوسط معدوم، وانحراف $1/\sqrt{T}$، ومنه يعطى مجال الثقة للمعامل $\rho(k)$ بـ . $\rho(k) = 0 \pm \frac{t_{\alpha/2}}{\sqrt{T}}$. إذا كان المعامل المحسوب $\hat{\rho}(k)$ خارج هذا المجال، فهو يختلف معنويا عن الصفر بنسبة معنوية α .

في حالة ما إذا كانت بيانات السلسلة مستقرة فإن معاملات الارتباط غالبا ما يكـون لها توزيع طبيعي متوسطه الحسابي 0 وتباينه $\frac{1}{T}$ ومن تم فإن حدود فترة الثقة عند مستوى معنوية 5% لعينة كبيرة الحجم هي $\pm 1.96\sqrt{1/T}$ وبالتالي إذا كان يقع خارج هذه الحدود فإننا نرفض فرضية العدم ويكون $\rho(k)$ مختلفا جوهريا عن الصفر.

ولإجراء اختبار مشترك لمعنوية معاملات الارتباط الذاتي كمجموعة نستخدم إحصائية Box-Pierce:

$$Q = T\sum_{k=1}^{K}\hat{\rho}^2(k)$$

التي تتوزع توزيع χ^2 بدرجة حرية K *و نسبة معنوية* α.

- ✓ إذا كان $Q > \chi_{\alpha}^2(K)$ نرفض فرضية العدم القائلة بأن كل معاملات الارتبـاط الذاتي مساوية للصفر وهذا يعني أن السلسلة غير مستقرة.
- ✓ إذا كان $Q < \chi_{\alpha}^2(K)$ نرفض الفرضية البديلة ونقبل فرضية العدم وهذا يعني أن السلسلة مستقرة (ساكنة).

كما أنه توجد إحصائية أخرى بديلة تستخدم في إجراء نفس الاختبار السابق تسمى بـ . Ljung-Box statistic وهي إحصائية Box-Pierce المعدلة و التي تعطى بالعلاقة التالية:

$$Q^* = T(T+2)\sum_{k=1}^{K}\frac{\hat{\rho}^2(k)}{T-k}$$

التي تتوزع توزيع χ^2 بدرجة حرية K *و نسبة معنوية* α. ويمكن استخدامها في حالـة العينات الصغيرة الحجم لأنهما تعطي نتائج أفضل من Q، مع كونها تصلح للعينات كبيرة الحجم[1].

1- عبد القادر محمد عبد القادر عطية، ص 620.

2.2. أهم اختبارات الجذر الوحدوي Unit Root tests:

إن اختبارات الجذر الوحدوي لا تعمل فقط على كشف مركبة الاتجاه العام، بل إنها تساعد على تحديد الطريقة المناسبة لجعل السلسلة مستقرة، ومن أجل فهم هذه الاختبارات لا بد من التفريق بين نوعين من النماذج غير المستقرة:

- ❖ النموذج Trend Stationary TS: هذه النماذج غير مستقرة، وتبرز عدم استقرارية تحديديه deterministic، وتأخذ الشكل $Y_t = f(t) + \varepsilon_t$ حيث $f(t)$ دالة كثير حدود للزمن (خطية أو غير خطية)، و ε_t تشويش أبيض، وأكثر هذه النماذج انتشارا يأخذ شكل كثير الحدود من الدرجة الأولى، ويكتب من الشكل $Y_t = a_0 + a_1\,t + \varepsilon_t$. هذا النموذج غير مستقر، لأن متوسطه $E(Y_t)$ مرتبط بالزمن، لكننا نجعله مستقرا بتقدير المعالم $\hat{a}_1, \hat{a}_0$ بطريقة المربعات الصغرى العادية، وطرح المقدار $\hat{a}_0 + \hat{a}_1\,t$ من Y_t، أي: $Y_t - \hat{a}_0 + \hat{a}_1\,t$.
- ❖ النموذج Differency Stationary DS: هذه النماذج أيضا غير مستقرة وتبرز عدم استقرارية عشوائية Stochastic، وتأخذ الشكل $Y_t = Y_{t-1} + \beta + \varepsilon_t$ ويمكننا جعلها مستقرة باستعمال الفروقات أي: $\nabla^d Y_t = \beta + \varepsilon_t$ حيث: β ثابت حقيقي، و d: درجة الفروقات. وغالبا تُستعمل الفروقات من الدرجة الأولى في هذه النماذج $d = 1$، وتكتب من الشكل $\nabla Y_t = \beta + \varepsilon_t$ وتأخذ هذه النماذج شكلين:

1. إذا كانت $\beta = 0$: يسمى النموذج DS بدون مشتقة، ويكتب من الشكل: $Y_t = Y_{t-1} + \varepsilon_t$ وبما أن ε_t تشويش أبيض، فإن النموذج يسمى " نموذج السير العشوائي Random Walk Model" وهو كثير الاستعمال في دراسة الأسواق المالية.
2. إذا كانت $\beta \neq 0$: يسمى النموذج DS بالمشتقة، ويكتب من الشكل $Y_t = Y_{t-1} + \beta + \varepsilon_t$.

1.2.2. اختبار ديكي– فولر Dickey-Fuller (DF) test :

تعمل اختبارات ديكي– فولار (Dickey-Fuller, 1979) على البحث في الاستقرارية أو عدمها لسلسلة زمنية ما، وذلك بتحديد مركبة الاتجاه العام، سواء كانت تحديدية (deterministic) أو عشوائية (Stochastic). لعرض هذا الاختبار نبدأ بنموذج السير العشوائي التالي الذي يسمى بنموذج الانحدار الذاتي من الدرجة الأولى AR(1)، والذي يكتب على الشكل:

$$Y_t = Y_{t-1} + \varepsilon_t$$

حيث ε_t: حد الخطأ العشوائي، والذي يُفترض فيه: وسط حسابي معدوم، تباين ثابت، وقيم غير مرتبطة (عندئذ يسمى حد الخطأ أو التشويش الأبيض).

يلاحظ أن معامل الانحدار يساوي الواحد 1، وإذا كان هذا هو الأمر في الواقع، فإن هذا يؤدي إلى وجود مشكلة الجذر الوحدوي الذي يعني عدم استقرار بيانات السلسلة، حيث يوجد هناك اتجاه في البيانات. لذا إذا قمنا بتقدير الصيغة التالية: $Y_t = \phi_1 Y_{t-1} + \varepsilon_t$، واتضح أن $\hat{\phi} = 1$ فإن المتغير Y_t يكون له جذر وحدوي، ويعاني من مشكلة عدم الاستقرار. وتعرف السلسلة التي يوجد لها جذر مساو للوحدة (كما ذكرنا أعلاه) بسلسلة السير العشوائي(Random Walk Time Series وهي إحدى الأمثلة للسلسلة غير المستقرة.

و بطرح Y_{t-1} من طرفي المعادلة $Y_t = \phi Y_{t-1} + \varepsilon_t$ نتحصل على الصيغة التالية:

$$\nabla Y_t = (\phi - 1) Y_{t-1} + \varepsilon_t$$

$$\nabla Y_t = \lambda \, Y_{t-1} + \varepsilon_t \quad , \quad \phi - 1 = \lambda$$

– حيث: $\nabla Y_t = Y_t - Y_{t-1}$، والآن أصبحت الفرضيات من الشكل:

$$H_0 : \quad \lambda = 0$$

$$H_1 : \quad \lambda \neq 0$$

ويلاحظ أنه إذا ثَبُتَ في الواقع أن $\lambda = 0$، فإن $\nabla Y_t = \varepsilon_t$، وعندئذ يُقال أن سلسلة الفروقات من الدرجة الأولى من السير العشوائي مستقرة، ولذا فإن السلسلة الأصلية

تكون متكاملة من الرتبة الأولى Integrated of Order 1، ونرمز لها بـ . *I(1)*. أما إذا كانت السلسلة مستقرة بعد الحصول على الفروقات من الدرجة الثانية (الفروقات الأولى للفروقات الأولى)، فإن السلسلة الأصلية تكون متكاملة من الرتبة الثانية أي *I(2)*، وهكذا...

وإذا كانت السلسلة الأصلية مستقرة يقال أنها متكاملة من الرتبة صفر أي *I(0)*.

ولاختبار مدى استقرار السلسلة نتبع الخطوات التالية:

1. نقوم بحساب ما يسمى بـ τ. بعد تقدير الصيغة $Y_t = \phi Y_{t-1} + \varepsilon_t$، بقسمة $\hat{\phi}$ على الخطأ المعياري لها، أي: $\tau = \dfrac{\hat{\phi}}{\hat{\sigma}_{\hat{\phi}}}$.
2. لا نستطيع مقارنة τ المحسوبة بقيم t المجدولة، حتى في العينات الكبيرة، لأنها لا تتبع هذا التوزيع، وإنما نبحث عن τ الجدولية في جداول معدة خصيصا بواسطة Dickey & Fuller، ولذا يُعرف هذا الاختبار باختبار Dickey-(DF-test) Fuller Test.
3. القرار:

- إذا كانت τ_c المحسوبة > τ_t المجدولة: نرفض فرضية العدم $H_0: \phi = 1$ أو $\lambda = 0$، ونقبل الفرضية البديلة $H_1: \phi \neq 1$ (أو $\lambda \neq 0$)، وبالتالي تكون السلسلة مستقرة.
- إذا كانت τ_c المحسوبة < τ_t المجدولة: نقبل فرضية العدم $H_0: \phi = 1$ ونرفض الفرضية البديلة $H_1: \phi \neq 1$، وفي هذه الحالة تكون السلسلة غير مستقرة.

ولقد جرت العادة على إجراء اختبار Dickey-Fuller باستخدام عدد من صيغ الانحدار تتمثل في[1]:

1- عبد القادر محمد عبد القادر عطية، ص 623.

$$\nabla Y_t = (\phi - 1)Y_{t-1} + \varepsilon_t$$
$$\nabla Y_t = (\phi - 1)Y_{t-1} + c + \varepsilon_t$$
$$\nabla Y_t = (\phi - 1)Y_{t-1} + c + b\,t + \varepsilon_t$$

وإذا وضعنا $\lambda = \phi - 1$ تصبح:

$$\Delta Y_t = \lambda\, Y_{t-1} + \varepsilon_t \qquad ....(1)$$
$$\Delta Y_t = \lambda\, Y_{t-1} + c + \varepsilon_t \qquad ....(2)$$
$$\Delta Y_t = \lambda\, Y_{t-1} + c + b\,t + \varepsilon_t \qquad ....(3)$$

حيث أن اختبار الفرضية $H_0 : \lambda = 0$ هو نفسه اختبار الفرضية $H_0\ \phi = 1$. مع مراعاة أنه تم إدخال الحد الثابت c في الصيغة (2)، وإدخال حد للاتجاه العام يتمثل في الـ زمن t في الصيغة (3).

وفي كل صيغة من الصيغ الثلاثة تكون الفروضيات من الشكل:

$$H_0 : \lambda = 0 \quad (\phi = 1)$$
$$H_1 : \lambda \neq 0 \quad (\phi \neq 1)$$

إن مبدأ هذا الاختبار بسيط هو:

- إذا تحققت الفرضية $H_0 : \phi = 1$ ($H_0 : \lambda = 0$) في أحد النماذج الثلاثة فـ إن السلسلة غير مستقرة.
- في النموذج (3)، إذا قبلنا الفرضية البديلة ($H_1 : \phi \neq 1$)، وكانت b معنوي ـا مختلف عن الصفر، فإن النموذج من النوع TS ويرجع مستقرا بطريقة الانحدار كما بيناها سابقا.
- حسب الفرضية H_0، فإن القواعد الإحصائية الاعتيادية من غير الممكن تطبيقها من أجل الاختبار. لذلك عمد ديكي وفولار إلى دراسة التوزيع التقاربي للمقدر $\hat{\phi}$، وذلك بمساعدة محاكاة مونتي-كارلو Monte-Carlo simulation، حيث جدولوا القيم الحرجة من أجل عينات ذات أطوال مختلفة، هذه الجداول شبيهة بجداول ستيودنت. وفي حالة وجود مشكلة الارتباط الذاتي بالحد العشوائي ε_t فإن الصيغة الملائمة للاستخدام هي اختبار ديكي فولار المطور.

في النماذج السابقة عند استعمالنا لاختبار ديكي-فولار البسيط، فـ إن النمـوذج ε_t عبارة عن صدمات عشوائية افتراضا، ففي حالة وجود ارتباط ذاتي بين الأخطـاء طـور ديكي وفولار (1981) اختبارا يسمى باختبار ديكي فولار المطور Augmented Dickey-Fuller (ADF) test.

إن اختبارات ADF ترتكز على الفرضية $H_1: |\phi| < 1$، وعلى التقدير بواسطة المربعات الصغرى[1]:

$$\nabla Y_t = \lambda Y_{t-1} - \sum_{j=2}^{p} \phi_j \nabla Y_{t-j+1} + \varepsilon_t \qquad ....(4)$$

$$\nabla Y_t = \lambda Y_{t-1} - \sum_{j=2}^{p} \phi_j \nabla Y_{t-j+1} + c + \varepsilon_t \qquad ....(5)$$

$$\nabla Y_t = \lambda Y_{t-1} - \sum_{j=2}^{p} \phi_j \nabla Y_{t-j+1} + c + b\,t + \varepsilon_t \qquad ....(6)$$

- نستطيع أن نحدد القيمة p حسب معيار Akaike أو معيار Schwarz.

إن اختبار ADF يحمل نفس خصائص اختبار DF، بحيث يستخدم الفروقـات ذات الفجوة الزمنيـة ∇Y_{t-j+1}، حيـث $\nabla Y_{t-1} = Y_{t-1} - Y_{t-2}$، $\nabla Y_{t-2} = Y_{t-2} - Y_{t-3}$،...إلخ، ويتم إدراج عدد من الفروقات ذات الفجوة الزمنية حتى تختفي مشكلة الارتباط الذاتي[2].

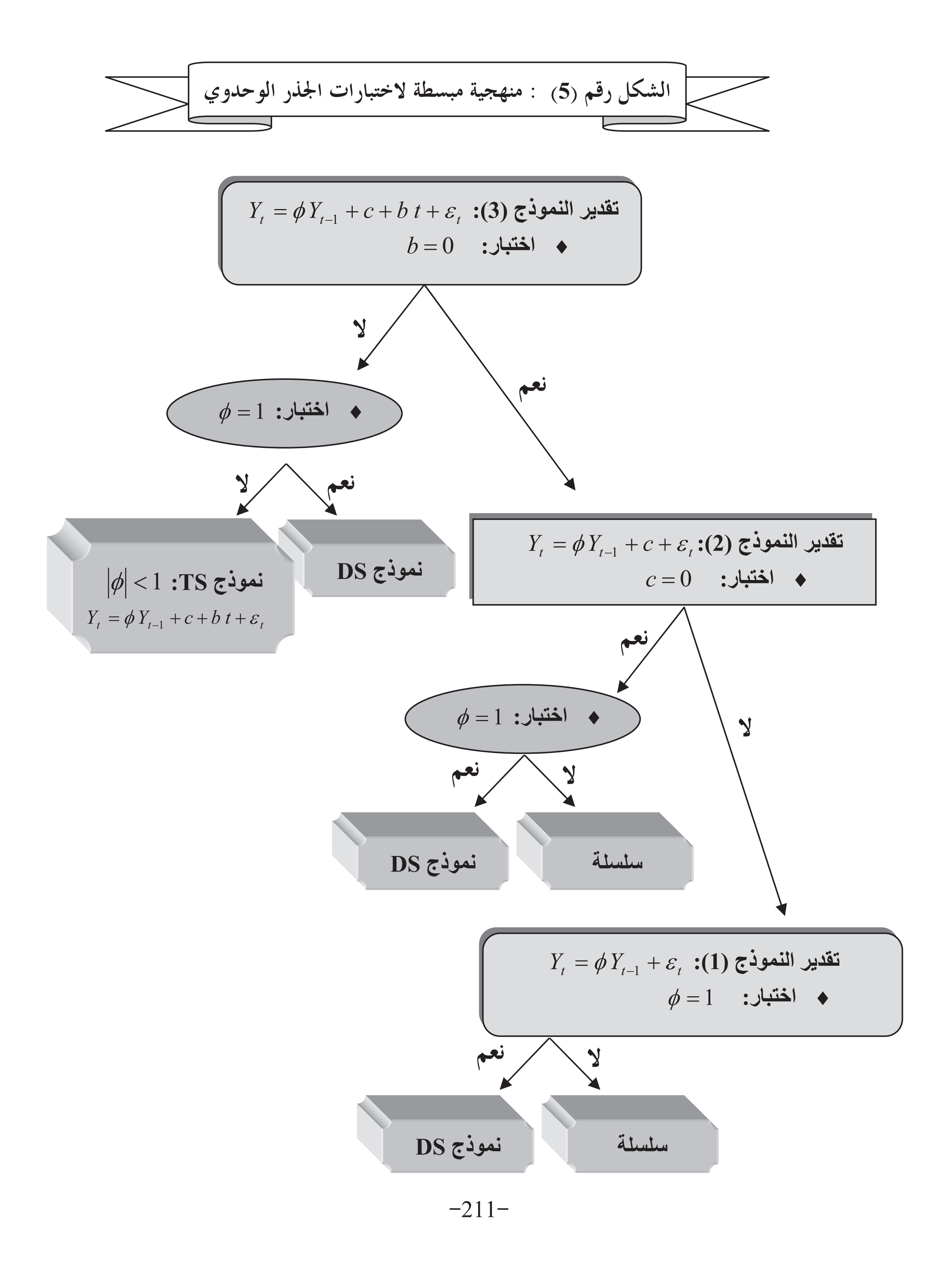
وفيما يلي صورة مبسطة لمنهجية اختبارات الجذر الوحدوي لـ ديكي- فولار[3]:

1- Bourbonnais (2003), p. 234.

2- عبد القادر محمد عبد القادر عطية، ص 623.

3- Bourbonnais (2003), p. 236

الشكل رقم (5) : منهجية مبسطة لاختبارات الجذر الوحدوي



2.2.2. اختبار فيليبس و بيرون Phillips and Perron test (1988)::

يعتبر هذا الاختبار غير المعلمي فعالا، حيث يأخذ بعين الاعتبـار التبـاين الشـرطي للأخطاء، فهو يسمح بإلغاء التحيزات الناتجة عن المميزات الخاصة للتذبذبات العشـوائية، حيث اعتمد Philips and Perron (1988) نفس التوزيعـات المحـدودة لاختبـاري DFوADF[1]. ويجرى هذا الاختبار في أربعة مراحل:

1. تقدير بواسطة OLS النماذج الثلاثة القاعدية لاختبـار Dickey-Fuller، مـع حساب الإحصائيات المرافقة.
2. تقدير التباين قصير المدى: $\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t^2$، حيث $\hat{\varepsilon}_t$ تمثل البواقي.
3. تقدير المعامل المصحح s_1^2، المُسمَّى التباين طويل المدى، والمستخرج من خـلال التباينات المشتركة لبواقي النماذج السابقة، حيث:

$$s_1^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t^2 + 2\sum_{i=1}^{l}\left(1-\frac{i}{l+1}\right)\frac{1}{T}\sum_{t=i+1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_{t-i}$$

من أجل تقدير هذا التباين يجب من الضروري إيجاد عـدد التبـاطؤات l Newey-West، المقدر بدلالة عدد المشاهدات الكلية T، على النحو التالي: $l \approx 4\left(\frac{T}{100}\right)^{2/9}$

4. حساب إحصائية فيليبس وبـيرون: $t_{\hat{\phi}}^* = \sqrt{k}\times\frac{(\hat{\phi}-1)}{\hat{\sigma}_{\hat{\phi}}} + \frac{T(k-1)\hat{\sigma}_{\hat{\phi}}}{\sqrt{k}}$ مـع $k = \frac{\hat{\sigma}^2}{s_1^2}$، والذي يساوي 1 – في الحالة التقاربية (asymptotic)– عندما تكون $\hat{\varepsilon}_t$ تشويشا أبيض. هذه الإحصائية تقارن مع القيمة الحرجة لجدول ماك كينون MacKinnon.

1- قبلي زهير، 1999، ص 50.

3.2.2. اختبار KPSS (KPSS Test, 1992)[1]:

اقترح Kwiatkowski.; Phillips.; Schmidt; Shin (1992) استخدام اختبار مضاعف لاغرانج، لاختبار فرضية العدم التي تقرر الاستقرارية للسلسلة. ويكون اختبار KPSS على المراحل التالية:

1. فبعد تقدير النماذج (2) أو (3)، نحسب المجموع الجزئي للبواقي: $S_t = \sum_{i=1}^{t} \hat{\varepsilon}_i$
2. نقدر التباين الطويل الأجل s_1^2 بنفس طريقة اختبار فليبس وبيرون.
3. نحسب إحصائية اختبار KPSS من العلاقة: $LM = \frac{1}{s_1^2} \frac{\sum_{t=1}^{T} S_t^2}{T^2}$

- نرفض فرضية العدم (فرضية الاستقرار): إذا كانت الإحصائية المحسوبة LM أكبر من القيمة الحرجة المستخرجة من الجدول المعد من طرف Kwiatkowski, Phillips, Schmidt.and Shin (1992.).
- نقبل بفرضية الاستقرار: إذا كانت الإحصائية LM أصغر من القيمة الحرجة.

إلا أن هناك اختبارات أخرى لاختبار الجدر الوحدوي منها اختبار HEGY[2] الذي يختبر وجود مركبة موسمية في السلسلة الزمنية. اقترح هذا الاختبار Hyllebeg and al (1990). هذه الإحصائية تقارن مع القيم الحرجة لـ Franses and Hobijn (1997).

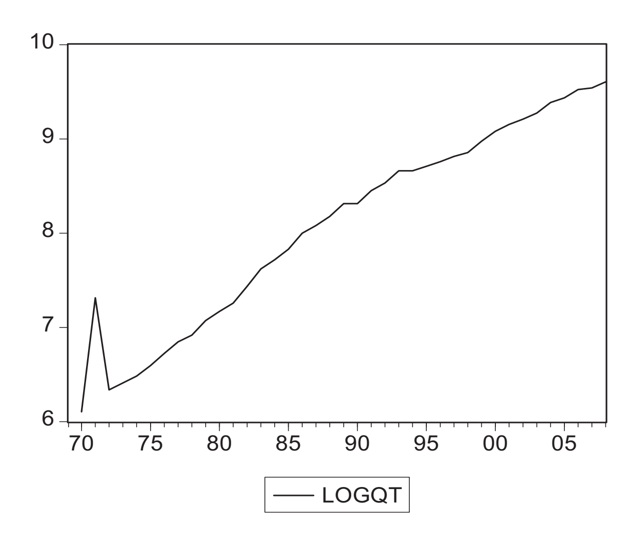
مثال 1:

الشكل البياني (6) يوضح تطور الطلب على الكهرباء في الجزائر (باللوغاريتم) خلال الفترة 1970 و 2008 على شكل معطيات سنوية:

1- Kwiatkowski, Phillips, Schmidt.and Shin (1992.)

2- Hylleberg, Engle, Granger and Yoo (1990). pp. 215-238.

الشكل رقم (6): تطور الطلب على الكهرباء في الجزائر باللوغاريتم (1970-2008)



الهدف الرئيسي هو معرفة طبيعة التغيرات التي تطرأ على قيم الطلب على الكهرباء في الجزائر في الفترات الزمنية من أجل استخراج في الأخير القيم المتوقعة لهذه الظاهرة. للقيام بعملية النمذجة، فلا بد أولا من دراسة استقرارية السلسلة. نلاحظ من خلال الشكل (6) أن هذه السلسلة تحتوي على اتجاه عام لأنها لا تتذبذب حول وسط حسابي ثابت. تكون السلسلة مستقرة، إذا كانت معاملات دالة الارتباط الذاتي $\rho(k)$ لا تختلف معنويا ع ن الصفر من أجل كل $k > 0$. والشكل التالي يبين دالة الارتباط الذاتي البسيطة والجزئي ة للسلسلة محل الدراسة:

الشكل رقم (7): دالة الارتباط الذاتي و الجزئي لسلسلة الطلب على الكهرباء باللوغاريتم

Sample: 1969 2008
Included observations: 39

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.893	0.893	33.539	0.000
		2	0.867	0.347	66.057	0.000
		3	0.789	-0.148	93.707	0.000
		4	0.710	-0.194	116.75	0.000
		5	0.630	-0.082	135.39	0.000
		6	0.551	-0.008	150.11	0.000
		7	0.473	-0.018	161.29	0.000
		8	0.398	-0.034	169.47	0.000
		9	0.322	-0.063	174.98	0.000
		10	0.250	-0.046	178.42	0.000
		11	0.183	-0.011	180.34	0.000
		12	0.117	-0.035	181.15	0.000
		13	0.054	-0.054	181.32	0.000
		14	-0.004	-0.032	181.32	0.000
		15	-0.058	-0.022	181.55	0.000
		16	-0.115	-0.065	182.46	0.000

نلاحظ من خلال دالة الارتباط الذاتي، أن المعاملات المحسوبة من أجل الفجوات $k=1,....,9$ تختلف معنويا عن الصفر عند مستوى معنوية 5% (خارج مجال الثقة $\left[\frac{-1.96}{\sqrt{T}},\frac{+1.96}{\sqrt{T}}\right]$)، أي تتناقص بوتيرة بطيئة نحو الصفر ولإثبات هذا نستعمل اختبار Ljung-Box لدراسة المعنوية الكلية لمعاملات دالة الارتباط الذاتي ذات الفجوات $k \leq 16$ أعلاه، حيث توافق إحصائية الاختبار المحسوبة Q^* آخر قيمة في العمود Q-Stat في الشكل أعلاه ، أي:

$$Q^* = T(T+2)\sum_{k=1}^{16}\frac{\hat{\rho}^2(k)}{T-k} = 39(39+2)\sum_{k=1}^{16}\frac{\hat{\rho}^2(k)}{39-k} = 182.46 > \chi^2_{0.05}(16) = 26.30$$

لدينا الإحصائية المحسوبة $Q^* = 182.46$ أكبر من الإحصائية المجدولة $\chi^2_{0.05}(16) = 26.30$ ومنه نرفض فرضية العدم القائل بأن كل معاملات الارتباط الذاتي تساوي معنويا الصفر عند مستوى معنوية 5%.

للكشف عن مركبة الاتجاه العام و تحديد الطريقة المناسبة لجعل سلسلة الطلب على الكهرباء مستقرة، نقوم باستعمال اختبارات الجذر الوحدوي. بالاستعانة ببرمجية Eviews، يمكن اختبار عدم استقرارية السلسلة وفق منهجية Dickey-Fuller و Philips-Perronو KPSS.

باستعمال برمجية Eviews، نتحصل على نتائج تقدير معالم النموذج (3) بطريقة المربعات الصغرى كما يلي:

$$\Delta Y_t = \text{-}0.812 Y_{t-1} + 6.665 + 0.072\, t + \hat{\varepsilon}_t$$
$$(\text{-}4.948) \quad (5.015) \; (4.725)$$

حيث أن القيم التي بين قوسين (.) هي قيم ستيودنت.

من الملاحظ من نتائج التقدير أن معامل الاتجاه العام b يختلف معنويا عن الصفر بنسبة معنوية 5%، نرفض الفرضية H_0 و للمعامل ϕ معنوية إحصائية، فهو يختلف معنويا عن الواحد أي أن λ (معامل Y_{t-1}) يختلف معنويا عن الصفر بنسبة معنوية 5%. نستنتج أنه يجب قبول فرضية TS أي أن عدم استقرار السيرورة التي تشرح الطلب على الكهرباء في الجزائر ناجمة عن وجود اتجاه عام ثابت و ليس عشوائيا. بعد تقدير المعادلة (1)، نحصل على النتائج التالية:

$$\Delta Y_t = 0.010 Y_{t-1} + \hat{\varepsilon}_t$$
$$(0.614)$$

نلاحظ أن القيمة المحسوبة $t_c = 0.614 \; > \; t_{Tab} = -1.94$ بالقيمة المطلقة، نقبل الفرضية H_0 أي أن السلسلة غير مستقرة.

الجدول التالي يعطي نتائج اختبارات الجذر الوحدوي (ADF، PP، KPSS) :

الجدول (1): نتائج اختبارات الجذر الوحدوي

نوع الاختبار	نوع النموذج	القيمة المحسوبة	القيمة الحرجة
اختبار ADF جذر وحدوي: H_0	النموذج (4)	1.751487	-1.9510
	النموذج (5)	-2.087770	-2.9499
	النموذج (6)	-1.213851	-3.5468
اختبار Philips-Perron (PP) جذر وحدوي: H_0	النموذج (1)	1.343037	-1.9498
	النموذج (2)	-1.347534	-2.9399
	النموذج (3)	-3.454286	-3.5312
اختبار KPSS استقرارية: H_0	النموذج (2)	0.6579	0.463
	النموذج (3)	0.1697	0.146

بالنسبة لاختبار ADF، عدد التباطؤات التي من خلالها معيار AIC أصغر ما يمكن هو 4، أما بالنسبة لاختباري PP و KPSS، عدد التباطؤات l Newey-West يساوي 6. من خلال هذه النتائج نستنتج أن سلسلة الطلب على الكهرباء غير مستقرة وتحتوي على جذر وحدوي، لإزالة مركبة الاتجاه العام، فلا بد من استعمال تقنية المربعات الصغرى و لكن يمكن أيضا القيام بحساب الفروقات من الدرجة الأولى $d=1$. الشكل التالي يبين دالة الارتباط الذاتي البسيطة والجزئية للسلسلة ذات الفروقات من الدرجة الأولى:

الشكل رقم (8): دالة الارتباط الذاتي و الجزئي للسلسلة ذات الفروقات من الدرجة الأولى

Included observations: 38

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.472	-0.472	9.1674	0.002
		2	0.003	-0.283	9.1678	0.010
		3	-0.009	-0.196	9.1709	0.027
		4	0.000	-0.141	9.1709	0.057
		5	0.003	-0.098	9.1715	0.102
		6	0.028	-0.027	9.2085	0.162
		7	-0.030	-0.034	9.2528	0.235
		8	0.024	0.001	9.2815	0.319
		9	0.005	0.027	9.2829	0.412
		10	-0.038	-0.021	9.3625	0.498
		11	-0.002	-0.047	9.3628	0.588
		12	0.034	-0.005	9.4319	0.666
		13	-0.006	0.008	9.4338	0.739
		14	-0.030	-0.033	9.4900	0.798
		15	0.036	0.008	9.5779	0.845
		16	-0.009	0.016	9.5842	0.887

نلاحظ من الشكل (8) أن المعاملات المحسوبة من أجل الفجوات $k = 2,....,16$ تساوي معنويا الصفر (داخل مجال الثقة $\left[\frac{-1.96}{\sqrt{T}}, \frac{+1.96}{\sqrt{T}}\right]$)، أي تتناقص بوتيرة سريعة نحو الصفر ويمكن التأكد من ذلك باستعمال اختبار Ljung-Box، لدينا:

$$Q^* = T(T+2)\sum_{k=1}^{16}\frac{\hat{\rho}^2(k)}{T-k} = 38(38+2)\sum_{k=1}^{16}\frac{\hat{\rho}^2(k)}{38-k} = 9.58 < \chi^2_{0.05}(16) = 26.30$$

لدينا الإحصائية المحسوبة $Q^* = 9.58$ أصغر من الإحصائية المجدولة $\chi^2_{0.05}(16) = 26.30$ ومنه نقبل فرضية العدم، أي أن كل معاملات الارتباط الذاتي تساوي معنويا الصفر.

3. اختبارات التوزيع الطبيعي Normality Tests[1]:

للبدء بدراسة السلوك الدوري لأي سلسلة زمنية مستقرة، فلا بد أولا من دراسة التوزيع الاحتمالي الذي تخضع له أي ظاهرة من أجل إعطاء نظرة أولية حول طبيعة هذه

1- Jarque and Bera (1987)

السلسلة المستقرة. نذكّر أنه من صفات التوزيع الطبيعي ينبغي أن يكون معامل Skewness معدوما و معامل Kurtosis مساويا إلى 3. فالقانون الطبيعي يتميز بالتناظر بالنسبة إلى المتوسط و باحتمال ضعيف للقيم الشاذة. يعتمد اختبار Jarque و Bera على معاملي التفلطح Kurtosis و التناظر Skewness.

إذا كان العزم الممركز من الدرجة k للسلسلة Y_t من الشكل $\mu_k = \frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \overline{Y}\right)^k$

فإن معامل Skewness يكتب كما يلي:

$$S = \frac{\left[\frac{1}{T}\sum_{t=1}^{T}(Y_t - m)^3\right]^2}{\left[\frac{1}{T}\sum_{t=1}^{T}(Y_t - m)^2\right]^3} = \frac{\mu_3^2}{\mu_2^3} = \beta_1$$

أما معامل Kurtosis فهو:

$$K = \frac{\frac{1}{T}\sum_{t=1}^{T}(Y_t - m)^4}{\left[\frac{1}{T}\sum_{t=1}^{T}(Y_t - m)^2\right]^2} = \frac{\mu_4}{\mu_2^2} = \beta_2$$

حيث m المتوسط الحسابي للسلسلة الزمنية المستقرة. إذا كان التوزيع طبيعيا وعدد المشاهدات كبيرا $n>30$، فإن:

$$\beta_1^{1/2} \sim N\left(0, \sqrt{\frac{6}{T}}\right)$$

$$\beta_2 \sim N\left(3, \sqrt{\frac{24}{T}}\right)$$

كما أشرنا سابقا، فاختبار جارك و بيرا يجمع بين المعاملين السابقين، فإذا كانت $\beta_1^{1/2}, \beta_2$ تتبعان التوزيع الطبيعي، فإن القيمة S تتبع توزيع χ^2 بدرجات حرية 2 حيث: $JB = \frac{T}{6}\beta_1 + \frac{T}{24}\left(\beta_2 - 3\right)^2 \sim \chi_\alpha^2(2)$. يتم إذن اختبار الفرضية التالية:

$$H_0 : \beta_1^{1/2} = \beta_2 - 3 = 0$$

إذا كانت $JB > \chi_\alpha^2(2)$ فإننا نرفض فرضية التوزيع الطبيعي للسلسلة بنسبة معنوية α.

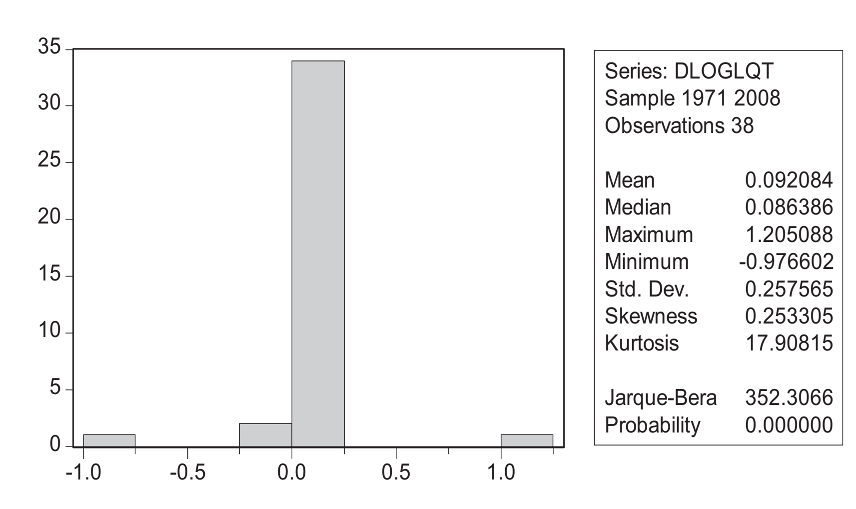
هناك اختبارات أخرى كثيرة تمكننا من دراسة طبيعة التوزيع الاحتمالي للسلسلة الزمنية المستقرة، نذكر منها على سبيل المثال طريقة النواة لتقدير دالة الكثافة التي تعتمد على معلم التمهيد الذي يتمثل في النافذة، فيتم في هذه الحالة مقارنة دالة الكثافة المقدرة بطريقة النواة بدالة كثافة التوزيع الطبيعي، فإذا كان هناك تقارب بينهما فإن التوزيع طبيعي[1].

تجدر الإشارة إلى أن السلاسل الزمنية المالية تتميز بتوزيع غير طبيعي حيث أن هذا التوزيع غير متناظر، فعدم التناظر يعتبر إشارة إلى عدم خطية النموذج الذي يتلاءم مع الظاهرة المدروسة باعتبار وجود تذبذبات غير ثابتة في التباين الشرطي للصدمات التي تطرأ على السوق المالي. سنتطرق إلى هذا النوع من المشاكل بالتفصيل في الفصل الثامن.

مثال 2:

سنختبر في هذه المثال ما إذا كانت سلسلة الطلب على الكهرباء المستقرة تحمل خصائص التوزيع الطبيعي أم لا، من أجل هذا يمكننا استعمال اختبار Jarque-Berra و ذلك بالاستعانة ببرمجية Eviews:

الشكل رقم (9): معاملات التوزيع الطبيعي



1- سيتم دراسة هذا النوع من الطرق بالتفصيل في الفصل 9.

إن دراسة التوزيع الطبيعي لهذه السلسلة تتم انطلاقا من قيمة معامل التناظر والـ تفلطح Skewness و Kurtosis على الترتيب:

- اختبار Skewness (اختبار فرضية التنـ ـاظر): $H_0 : \nu_1 = 0$، نقـ ـوم بحسـ ـاب الإحصائية:

$$\nu_1 = \frac{\beta_1^{1/2} - 0}{\sqrt{\frac{6}{T}}} = \frac{0.2533 - 0}{\sqrt{\frac{6}{38}}} = 0.638 < 1.96$$

لدينا $\nu_1 < 1.96$ ومنه نقبل الفرضية $H_0 : \nu_1 = 0$ أي أن هذه السلسلة متناظرة.

- اختبار Kurtosis: (اختبار فرضية التفلطح الطبيعي) $H_0 : \nu_2 = 0$:

$$\nu_2 = \frac{\beta_2 - 3}{\sqrt{\frac{24}{T}}} = \frac{17.9081 - 3}{\sqrt{\frac{24}{38}}} = 18.75 > 1.96$$

بما أن $\nu_2 > 1.96$: نرفض فرضية التفلطح الطبيعي للسلسلة.

يمكن التأكد من ذلك باستعمال إحصائية Jarque-Bera، حيث نلاحـ ـظ أن هـ ـذه الأخيرة $JB = 352.306 > \chi^2_{0.05}(2) = 5.99$ وعليه السلسلة المستقرة لا تتوزع توزيعـ ـا طبيعيا...

4. اختبارات الاستقلالية Independence Tests:

1.4. اختبار Mizrach - Non parametric MizrachTest[1]:

طور Mizrach (1996) اختبارا غير معلمي يسمح باختبار فرضية استقلالية المشاهدات *independently and identically distributed* (*iid*) الخاصة بالسلاسل الزمنيـ ـة ضـ ـد فرضية بديلة غير معروفة، هذا الاختبار لا يكشف فقط وجود ارتباط عادي في السلاسل الزمنية، وإنما يكشف أيضا وجود بنية غير خطية.

1- Mizrach (1996).

ليكن $p_1, p_2, ..., p_{m-1}$ متتالية متزايدة لأعداد صحيحة على $[1, L]$، $L < T - m + 1$.
في حالة وجود ارتباط، لدينا:

$$P[y_{t+p_{m-1}} < \varepsilon, ..., y_{t+p_1} < \varepsilon, y_t < \varepsilon] = (P[y_t < \varepsilon])^m$$

نقدر التوزيعات المشتركة $F(y_t^m)$، و الهامشية $F(y_t)$ في هذه المعادلة بطريقة النواة $K : R \rightarrow R$:

$$K(y_t) = I(y_t < \varepsilon) = \begin{cases} 1, si : y_t > \varepsilon \\ 0, otherwise \end{cases} \equiv I(y_t, \varepsilon)$$

الاحتمال غير الشرطي المشترك من أجل القيم y أصغر من ε يعطى بالعلاقة:

$$\theta(m, \varepsilon) = \int_Y \prod_{i=0}^{m-1} I(y_{t+p_i}, \varepsilon) dF(y_t)$$

إذن الإحصائية U U-statistic يمكن كتابتها على الشكل:

$$\theta(k, n, \varepsilon) = \sum_{t=1}^{n} \prod_{i=0}^{m-1} I(y_{t+p_i}, \varepsilon) / n \quad , \qquad n = T - m + 1$$

وكنتيجة لذلك، يمكن بناء الاختبار اللامعلمي لـ . . Mizrach و ذلك باستعمال المقدرات المتسقة للعزمين الأول والثاني للإحصائية U. تحت ظل قبول فرضية العدم لاستقلالية المشاهدات، إحصائية Mizrach تتبع القانون الطبيعي الممركز والمختصر[1]. من أجل: $p_i \in [1, L], i = 1, ..., m-1, L < n$، إذن إذا كان $\theta(k, \varepsilon) > 0$، فإن:

$$\sqrt{n} \frac{[\theta(m, n, \varepsilon) - \theta(m-1, n, \varepsilon)\theta(1, n, \varepsilon)]}{[\theta(m-1, n, \varepsilon)\theta(1, n, \varepsilon)(1 - \theta(m-1, n, \varepsilon))(1 - \theta(1, n, \varepsilon))]^{1/2}} \sim N(0,1)$$

إذا كانت إحصائية Mizrach أكبر من القيمة الحرجة للتوزيع الطبيعي عند مستوى معنوية α، فإننا نرفض فرضية العدم H_0، ومنه تكون السلسلة ذات بنية ارتباط. هناك ربط بين رفض الاستقلالية و قابلية السلسلة للتنبؤ، فإذا رفضنا H_0، فهذا يعني أن السلسلة تتميز ببنية ارتباط أي أن الظاهرة الاقتصادية قابلة للتنبؤ على المدى القصير أي أن التقلبات الدورية لهذه الظاهرة ما هي إلا نتيجة صدمات خارجية عابرة.

1- Mizrach (1996, p.7)

ويعتبر اختبار Mizrach الأكثر فعالية وقوة من بـ ـين الاختبـ ـارات غـ ـير المعلمية للاستقلالية، حيث يستطيع أن يكشف لنا عن كل أنواع الارتباط، ويساعدنا في تحديد أحسن سيرورة معممة للمعطيات Data generator Process (DGP).

2.4. اختبار Non parametric BDS Test –BDS[1]:

اقترح Brock, Dechert and Scheinkman (1987) اختبارا غير معلمي يعتمد على تكامل الارتباط لـ Grassbege et Procaccia. يعتبر هذا أكثر قوة من اختبار Mizrach عندما يكون حجم العينة يفوق 1000 مشاهدة. نختبر الفرضية القائلة بأن السلسلة مستقلة ومتماثلة التوزيع *independently and identically distributed* (*iid*) ضد فرضية الارتباط الخطي أو غير الخطي.

نذكر أن تكامل الارتباط يعرف كما يلي:

$$C(\varepsilon,k)=\frac{1}{n(n-1)}\sum_{i,j=1}^{n}H\left(\varepsilon-\left\|Y_i^k-Y_j^k\right\|\right)$$

حيث $n=T-k+1$ و $\left\|Y_i^k\right\|=\max\left|Y_i\right|$ و H هي دالة Heaviside:

$$H(y)=\begin{cases}1, & si: y>0\\ 0, & \text{otherwise}\end{cases}$$

بين Brock, Dechert and Scheinkman (1987) تحت فرضية *iid* أنه إذا كان $\sigma_k^2>0$، فإن:

$$T^{1/2}\left[C(\varepsilon,m,T)-(C(\varepsilon,m,T))^m\right]\rightarrow N(0,\sigma_m^2) \quad \text{مع} \quad T\rightarrow\infty$$

حيث:

$$\sigma_m^2=4\left[K^m+2\sum_{i=1}^{m-1}K^{m-i}C^{2i}+(k-1)^2C^{2m}-k^2KC^{2m-2}\right]$$

مع:

$$C=C(\varepsilon)=E\left(I(Y_i,Y_j;\varepsilon)\right)$$
$$K=K(\varepsilon)=E\left(I(Y_i,Y_j;\varepsilon)I(Y_j,Y_m;\varepsilon)\right)$$

و $I(a,b;\varepsilon)$ دالة Heaviside. أما $C(\varepsilon)$ مقدرة بـ $\hat{C}(\varepsilon,T)$ و $K(\varepsilon)$ بالمعادلة:

1- Brock, Dechert, Scheinkman (1987) and Brock, Dechert, Scheinkman and LeBaron (1996).

$$\hat{K}(\varepsilon,T)=\frac{6}{n(n-1)(n-2)}\sum_{i<j<k}I(Y_i^m,Y_j^m,Y_k^m)$$

و:

$$I(a,b,c)=\frac{1}{3}\left[I(a,b;\varepsilon)i(b,c;\varepsilon)+I(a,c;\varepsilon)I(c,b;\varepsilon)+I(b,a;\varepsilon)I(a,c;\varepsilon)\right]$$

إحصائية BDS معطاة بالعلاقة:

$$W(\varepsilon,m,T)=T^{1/2}D(\varepsilon,m,T)/\sigma_m(\varepsilon,T)$$

حيث:

$$D(\varepsilon,m,T)=C(\varepsilon,m,T)-(C(\varepsilon,m,T))^m$$

تنعدم هذه الإحصائية من أجل حجم عينة يؤول إلى ما لا نهاية إذا كانت السلسلة *iid* [1] وغير معدومة إذا كانت السيرورة تتميز بارتباط قوي. بالأخذ بعين الاعتبار أن $C(\varepsilon,1)\xrightarrow[T\to\infty]{}[C(\varepsilon,1)]^m$، يمكن كتابة المعادلة الأخيرة كما يلي:

$$W(\varepsilon,m)=T^{1/2}\frac{\left[C(\varepsilon,m)-(C(\varepsilon,1))^m\right]}{\sigma_m(\varepsilon)}$$

تحت ظل قبول فرضية السير العشوائي، تتوزع هذه الإحصائية توزيعا طبيعيا مركزا مختزلا. يتبين لنا أن W هي دالة لمجهولين: البعد *Embedding Dimension m* و *النواة* ε. يوجد علاقة مهمة تربط بين اختيار m و ε و خصائص العينة الصغيرة لإحصائية BDS. من أجل كل قيمة m، لا يجب أن يكون ε لا كبيرا و لا صغيرا. يتم إذن اختيار ε بحيث $\frac{1}{2}<\frac{\varepsilon}{\sigma}<2$ حيث σ هي الانحراف المعياري للسلسلة المدروسة. يرتبط اختيار البعد m بعدد المعطيات المتوفرة لدينا. التوزيع صحيح على عينة محدودة إذا كان $\frac{T}{m}>200$. بصفة عامة، تختبر إحصائية BDS فرضية العدم لسلسلة *iid*، فرفض هذه الفرضية يمكن أن يكون ناجما عن وجود بنية ارتباط في سيرورة عشوائية خطية أو بنية ارتباط غير خطي (عشوائي بحت أو مشوش Chaotic). يمكن القول أن هذا الاختبار يختبر أيضا قابلية السلسلة الزمنية للتنبؤ على المدى القصير أي يدرس طبيعة الصدمات الخارجية التي تطرأ على الأسواق المالية، حيث يعتبر هذا الاختبار أكثر شيوعا في دراسة السلاسل المالية.

1- Brock et Baek (1991).

مثال 3:

بالاستعانة بمعطيات المثال السابق، نختبر ما إذا كانت السلسلة تتميز ببنية ارتباط و بتوزيع متماثل و مستقل *iid*. للقيام بذلك، يمكن استخدام برمجية Eviews 5.0 و خوارزمية خاصة لحساب إحصائيات Mizrach. نتائج هذه الاختبارات مبينة في الجدول التالي:

الجدول (2): نتائج اختبارات BDS و Mizrach للاستقلالية

إحصائيات Mizrach	إحصائيات BDS		m
	Standard Deviation	Fraction of pairs	
$T \leq 1000$	6.7977	3.1623	2
3.3802	6.4106	2.2273	3
2.8870	6.0363	2.3775	4
2.7684	5.7549	2.4705	5
2.4346	5.5407	2.5269	6
2.2964	5.3694	2.0781	7
1.9861	5.2236	1.9793	8

من خلال هذه النتائج يتضح جليا أن السلسلة المستقرة للطلب على الكهرباء تتميز بارتباط قوي حيث أننا نرفض فرضية الاستقلالية *iid*، باعتبار أن من أجل $m = 2,3,...,8$ إحصائيات Mizrach و BDS أكبر تماما من القيمة المجدولة للتوزيع الطبيعي 1.96 عند مستوى معنوية 5%. يمكن القول أن سلسلة الطلب على الكهرباء قابلة للتنبؤ على المدى القصير وتعتبر طبيعة الصدمات في هذه الحالة عابرة..

5. النماذج الخطية للسلاسل الزمنية

هدفنا في هذا الجزء هو تقديم نماذج تشرح سلوك السلسلة الزمنية Y_t، هذه الأخيرة نشرحها بواسطة قيمها الحالية والماضية (المبطأة)، نبدأ تحليلنا ببناء نماذج مبسطة للسلاسل الزمنية من نوع المتوسط المتحرك(Moving Average (MA ، ونموذج الانحدار الذاتي Autoregressive (AR) بالنسبة للسيرورات المستقرة، فمثلا في نموذج المتوسط المتحرك

تكون السيرورة Y_t ممثلة بواسطة مجموع المرجَّحات للأخطاء العشوائية الحالية والمبطأة، أما في نموذج الانحدار الذاتي، فتعتمد السلسلة الزمنية Y_t على مجموع المرجحات لقيمها الماضية وحد الخطأ العشوائي، تشمل النماذج المختلطة النوعين المذكورين والتي تسمى بنماذج الانحدار الذاتي والمتوسط المتحرك (ARMA) Moving Average Models‒Autoregressive ، حيث تكون Y_t في نماذج (ARMA) عبارة عن دالة لكل من الأخطاء العشوائية و الظاهرة الاقتصادية الحالية والماضية.

1.5. نموذج المتوسط المتحرك (MA) Moving Average Models::

تكون كل ملاحظة من السلسلة الزمنية Y_t، في سيرورة المتوسط المتحرك من الدرجة $q \geq 1$ مُفسَّرة بواسطة متوسط مرجّح للأخطاء العشوائية التي نرمز لها بـ MA(q)، وتكتب معادلتها على الشكل:

$$Y_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q}$$

حيث أن $\theta_0, \theta_1, \theta_2, ..., \theta_q$ هي معالم النموذج التي يمكن أن تكون موجبة أو سالبة و $\varepsilon_t, \varepsilon_{t-1}, ..., \varepsilon_{t-q}$ متوسطات متحركة لقيم الحد العشوائي في الفترة t والفترات السابقة.

نفرض أن الأخطاء مُفسرة بواسطة سيرورة الاضطراب (التشويش) الأبيض، وكحالة خاصة هذه الأخطاء مستقلة ومتماثلة التوزيع $i.i.d$، إذا كانت: $E(\varepsilon_t) = 0$, $\text{var}(\varepsilon_t) = \sigma_\varepsilon^2$, $E(\varepsilon_t \varepsilon_{t-k}) = 0$ من أجل $k \neq 0$، فإن وسط السيرورة MA(q) يكون مُستقلا عن الزمن t ما دام $E(Y_t) = \theta_0$، ليصبح التباين المشترك لهذه السيرورة:

$$E(Y_t Y_{t-k}) = E\left[Y_{t-k}\left(\theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q}\right)\right]$$

أي:

$$\gamma(k) = E(\varepsilon_t \varepsilon_{t-k}) = 0, \quad k \neq 0$$

لتكن السيرورة MA(q)، الممثلة بالمتوسط θ_0 وتباين الأخطاء σ_ε^2. أما التباين $\gamma(0)$ لسيرورة المتوسط المتحرك $(k = 0)$، ذي الدرجة q فهو على الشكل:

$$\begin{aligned}\operatorname{var}(Y_t) &= \gamma(0) = E\left[(Y_t - \theta_0)^2\right] \\ &= E\left[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots.. + \theta_q\varepsilon_{t-q})(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots.. + \theta_q\varepsilon_{t-q})\right] \\ &= E\left[\varepsilon_t^2 + \theta_1^2\varepsilon_{t-1}^2 + \theta_2^2\varepsilon_{t-2}^2 + \ldots\ldots + \theta_q^2\varepsilon_{t-q}^2 + 2\theta_1\varepsilon_t\varepsilon_{t-1} + \ldots.\right] \\ &= \sigma_\varepsilon^2 + \theta_1^2\sigma_\varepsilon^2 + \theta_2^2\sigma_\varepsilon^2 + \ldots\ldots\ldots\ldots + \theta_q^2\sigma_\varepsilon^2 \\ &= \sigma_\varepsilon^2\left[1 + \theta_1^2 + \theta_2^2 + \ldots\ldots\ldots\ldots + \theta_q^2\right]\end{aligned}$$

$$\operatorname{var}(Y_t) = \gamma(0) = \sigma_\varepsilon^2\left[1 + \sum_{j=1}^{q}\theta_j^2\right]$$

وكمثال نأخذ سيرورة المتوسط المتحرك من الدرجة الأولى MA(1) :

$$Y_t = \theta_0 + \varepsilon_t + \theta_1\varepsilon_{t-1}$$

إن هذه السيرورة متوسطها θ_0، وتباينها $\operatorname{var}(Y_t) = \gamma(0) = \sigma_\varepsilon^2\left[1 + \theta_1^2\right]$، أما التباين المشترك فهو على الشكل:

$$\begin{aligned}\operatorname{cov}(Y_t, Y_{t-1}) &= \gamma(1) = E\left[(Y_t - \theta_0)(Y_{t-1} - \theta_0)\right] \\ &= E\left[(\varepsilon_t + \theta_1\varepsilon_{t-1})(\varepsilon_{t-1} + \theta_1\varepsilon_{t-2})\right] \\ &= \theta_1\sigma_\varepsilon^2\end{aligned}$$

وعلى العموم نحدد التباين المشترك لـ k فترة مبطأة على الشكل:

$$\begin{aligned}\operatorname{cov}(Y_t, Y_{t-k}) &= \gamma(k) = E\left[(Y_t - \theta_0)(Y_{t-k} - \theta_0)\right] \\ &= E\left[(\varepsilon_t + \theta_1\varepsilon_{t-1})(\varepsilon_{t-k} + \theta_1\varepsilon_{t-k-1})\right] = 0 \qquad : k > 1\end{aligned}$$

ومنه فإن للسيرورة MA(1) تباين مشترك معدوم لما يكون التباطؤ أكبر من فترة واحدة، أي أن كل قيمة للسلسلة الزمنية Y_t تكون مرتبطة مع Y_{t-1} و Y_{t+1} دون قيم السلسلة الزمنية الماضية والمستقبلية الأخرى Y_{t-k} و Y_{t+k}. وهذا يعني أن الحوادث الظاهرة في أكثر من فترة زمنية واحدة في الماضي ليس لها أثر على السيرورة حاليا، كما أن الذاكرة المحدودة لسيرورة المتوسط المتحرك توفر معلومة محدودة من أجل التنبؤ بنموذج المتوسط المتحرك في المستقبل، تكون هذه المعلومات مساوية لعدد فترات التباطؤ q، وفي مثالنا تكون فترة واحدة في المستقبل فقط.

ومنه نقول أن دالة الارتباط الذاتي للسيرورة MA(1) هي:

$$\rho(k)=\frac{\gamma(k)}{\gamma(0)}=\begin{cases}\dfrac{\theta_1}{\left(1+\theta_1^2\right)}\ , & k=1\\ 0\ , & k>1\end{cases}$$

يمكن القول أن لدالة الارتباط الذاتي $\rho(k)$ للسيرورة MA(q) q قيمـ ـة تختلـ ـف عـ ـن الصفر، وتساوي الصفر فقط لما يكون $k>q$، لذلك يتم الاعتماد على دالة الارتبـ ـاط الذاتي في تحديد درجة السيرورة MA(q).

2.5. نماذج الانحدار الذاتي(Autoregressive Models (AR :

طبقا لهذه النماذج تكون الملاحظة الحالية Y_t مُفسَّرة بواسـ ـطة متوسـ ـط التـ ـرجيح للملاحظات الماضية إلى فترة التأخير من الدرجة p مع الأخذ بعين الاعتبار حـ ـد الخط ـأ العشوائي في الفترة الحالية، ونسمي ذلك بنموذج الانحدار الذاتي للسلسلة الزمنيـ ـة Y_t ذي الدرجة p **(Autoregressive of order p)**.

يكتب نموذج الانحدار الذاتي من الدرجة p على الشكل:

$$Y_t=\phi_0+\phi_1 Y_{t-1}+\theta_2 Y_{t-2}+.....+\phi_p Y_{t-p}+\varepsilon_t$$

بمعنى:

$$Y_t=\phi_0+\sum_{i=1}^{p}\phi_i Y_{t-i}+\varepsilon_t$$

حيث Y_t: قيمة المتغير في الفترة الحالية t ، ε_t: حد الخطأ العشوائي في الفترة الحالية t ، $Y_{t-1},Y_{t-2},.......Y_{t-p}$: قيم المتغير في الفترات السابقة، ϕ_0: ثابت.

وعادة ما يكتب نموذج الانحدار الذاتي بواسطة معامل التأخير (أو التباطؤ) L :

$$Y_t=\phi_0+\phi_1 LY_t+\theta_2 L^2 Y_t+......+\phi_p L^p Y_t+\varepsilon_t$$

$$\Rightarrow\quad \left(1-\phi_1 L-\phi_2 L^2-.......-\phi_p L^p\right)Y_t=\phi_0+\varepsilon_t$$

$$\Rightarrow\quad \phi(L)Y_t=\phi_0+\varepsilon_t$$

حيث:

$$\phi(L)=1-\phi_1 L-\phi_2 L^2-.............-\phi_p L^p$$

إذا كانت السيرورة AR(p) أعلاه مستقرة، فإن وسطها الممثل بـ μ. ، يجب أن يكون غير متغير بالنسبة للزمن، أي:

$$E(Y_t) = E(Y_{t-1}) = E(Y_{t-2}) = ...\ ... = E(Y_{t-p}) = \mu$$

وهذا يعني:

$$E(Y_t) = \phi_0 + \phi_1 E(Y_{t-1}) + \phi_2 E(Y_{t-2}) + ...\ .... + \phi_p E(Y_{t-p}) + E(\varepsilon_t)$$

$$\mu = \phi_0 + \phi_1 \mu + \theta_2 \mu + ...\ .... + \phi_p \mu$$

$$\mu = \frac{\phi_0}{1 - \sum_{i=1}^{p} \phi_i}$$

إن العبارة الأخيرة والخاصة بمتوسط السيرورة AR(p) تُعطي لنا أيضا شرط الاستقرار، فإذا كان μ منتهيا فمن الضروري أن تكون $\sum_{i=1}^{p} \phi_i < 1$ ، إن هذا الشرط ضروري، لكنه غير كاف لضمان حالة الاستقرار، حيث هناك شروط أخرى يجب أن تتحقق.

وبوضع $y_t = Y_t - \phi_0$ ، وانطلاقا من نموذج الانحدار الذاتي المكتوب بواسطة معامل التأخير L يكون لدينا:

$$\phi(L) y_t = \varepsilon_t$$

ومنه فإن:

$$y_t = \phi^{-1}(L) \varepsilon_t$$

إذن لكي يكون النموذج AR(p) مستقرا يجب أن يكون قابلا للانعكاس (للقلب) invertible، أي يمكن كتابته على شكل نموذج نهائي للأخطاء العشوائية. وبعبارة أخرى يجب أن تكون جذور كثير الحدود $\phi(L)$ بالقيمة المطلقة أقل من الواحد[1].

لندرس الآن خصائص السيرورة AR(p) البسيطة بواسطة تحديد متوسطها، تباينها وتبايناتها المشتركة، ولنبدأ بالسيرورة AR(1) على الشكل $Y_t = \delta + \phi_1 Y_{t-1} + \varepsilon_t$.

إذا كانت $E(Y_t) = E(Y_{t-1}) = \mu$ فإن متوسط هذه السيرورة هو: $\mu = \frac{\phi_0}{1 - \phi_1}$

1- Bresson and Michaud (1995), p.22.

تكون السيرورة AR(1) أعلاه مستقرة إذا تحقق $|\phi_1|<1$.

لنحسب الآن تباين هذه السيرورة $\gamma(0)$، إذا وضعنا $\phi_0=0$ مع وجـ ـود الشـ ـرط $|\phi_1|<1$ يكون التباين ثابتا أي:

$$\operatorname{var}(Y_t)=\gamma(0)=E\left[(\phi_1 Y_{t-1}+\varepsilon_t)^2\right]$$
$$=\phi_1^2\gamma(0)+\sigma_\varepsilon^2$$

وبالترتيب نجد:

$$\operatorname{var}(Y_t)=\gamma(0)=\frac{\sigma_\varepsilon^2}{1-\phi_1^2}$$

أما التباينات المشتركة لـ Y_t. حول وسطها فهي:

$$\gamma(1)=E(Y_tY_{t-1})=\phi_1\gamma(0)$$
$$\gamma(2)=E(Y_tY_{t-2})=\phi_1^2\gamma(0)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\gamma(k)=E(Y_tY_{t-k})=\phi_1^k\gamma(0).$$

لتكن دالة الارتباط الذاتي للسيرورة وتنخفض هندسيا على الشكل:

$$\rho(k)=\frac{\gamma(k)}{\gamma(0)}=\begin{cases}1 & :k=0\\ \phi_1^k & :k=1,2,\ldots\end{cases}$$

حيث نلاحظ أن السيرورة AR(1) لها ذاكرة غير منتهية، وذلك لاعتماد القيمة الحاليـ ـة للسيرورة على كل القيم الماضية، بالرغم من أن تصرف هذه التبعية ينخفض مع الزمن.

إن أحد المشاكل المعروفة في بناء نماذج الانحدار الذاتي هي تحديد درجة السيرورة، فبالنسبة لنماذج المتوسط المتحرك يكون هذا المشكل بسيطا، حيث إذا كانت السيرورة من الدرجة q فإن الارتباطات الذاتية يجب أن تكون كلها قريبة من الصفر من أجل تباطؤات أكبر من q، وبالرغم من أن بعض المعلومات حول درجة الانحدار الذاتي يمكن الحصول عليها من السلوك الدوري لعينة دالة الارتباط الذاتي، فإن معلومات أكثر يمكن استنتاجها من دالة الارتباط الجزئية.

ولمعرفة هذه الأخيرة وكيفية استعمالها، نعتبر أولا التباينات المشتركة ودالة الارتباط الذاتي للسيرورة AR(p)، حيث نلاحظ أن التباين المشترك بتأخير k محدد من:

$$\gamma(k) = E\left[Y_{t-1}\left(\phi_1 Y_{t-1} + \theta_2 Y_{t-2} + ... \ ..... + \phi_p Y_{t-p} + \varepsilon_t\right)\right]$$

ولنترك الآن $k = 0,1,2,....., p$ لنحصل على $p+1$ معادلات فروقات[1] والتي يمك ـن حلها نهائيا من أجل: $\gamma(0), \gamma(1), ..., \gamma(p)$. حيث:

$$\begin{cases} \gamma(0) = \phi_1\gamma(1) + \phi_2\gamma(2) + ... + \phi_p\gamma(p) + \sigma_\varepsilon^2 \\ \gamma(1) = \phi_1\gamma(0) + \phi_2\gamma(1) + ... + \phi_p\gamma(p-1) \\ \gamma(2) = \phi_1\gamma(1) + \phi_2\gamma(0) + ... + \phi_p\gamma(p-2) \\ \vdots \\ \gamma(p) = \phi_1\gamma(p-1) + \phi_2\gamma(p-2) + ... + \phi_p\gamma(0) \end{cases}$$

وبالنسبة للتأخيرات $k > p$ تصبح لدينا:

$$\gamma(k) = \phi_1\gamma(k-1) + \phi_2\gamma(k-2) + ... + \phi_p\gamma(k-p) \quad : k > p$$

وللحصول على معادلات Yulle-Walker لدالة الارتباط الذاتي نقوم بقسمة التباين ـات المشتركة على التباين فنحصل على:

$$\begin{aligned} &\rho(1) = \phi_1 + \phi_2\rho(1) + ... + \phi_p\rho(p-1) \\ &\vdots \\ &\rho(p) = \phi_1\rho(p-1) + \phi_2\rho(p-2) + ... + \phi_p \end{aligned}$$

ومن أجل $k > p$ ينتج لدينا:

$$\rho(k) = \phi_1\rho(k-1) + \phi_2\rho(k-2) + ... + \phi_p\rho(k-p)$$

إذا كانت $\rho(1), \rho(2), ..., \rho(p)$ معروفة (مُقاسة من دالة الارتباط الذاتي المقدرة)، فإنه يمكن حل معادلات Yulle-Walker من أجل المعالم $\phi_i : i = 1,2,......p$، ولكن عمليا يتطلب حل هذه الأخيرة معرفة درجة الانحدار الذاتي p، وتحديد هذه الدرجة يعتبر أمرا

1- تسمى هذه المعادلة بمعادلة Yule-Walker.

صعبا، ولهذا نفترض أننا نحل معادلات Yulle-Walker من أجل القيم المثالية لـ p . أي نبدأ بوضع الفرضية $p=1$ ، ومن ثم يصبح لدينا $\rho(1)=\phi_1$ ، أو نستعمل الارتباطات الذاتية المقدرة $\hat{\rho}(1)=\hat{\phi}_1$ ، ولهذا إذا كانت لـ $\hat{\phi}_1$. معنوية إحصائية جيدة (مختلفة معنويا عن الصفر)، نقول أن سيرورة الانحدار الذاتي تكون على الأقل من الدرجة الأولى.

ونمثل تلك القيمة لـ $\hat{\phi}_1$. بواسطة $r(1)$، ثم نعتبر الفرضية $p=2$ ، أي AR(2) وللقيام بذلك نحل معادلات Yulle-Walker من أجل $p=2$. وهذا يعطي مجموعة جديدة من المقدرات $\hat{\phi}_2,\hat{\phi}_1$ ، حيث إذا كانت $\hat{\phi}_2$ لها معنوية إحصائية جيدة يمكن من الاستنتاج أن السيرورة على الأقل من الدرجة الثانية، بينما إذا كانت $\hat{\phi}_2$ قريبة من الصفر، نقول أن $p=1$ ، لنمثل قيمة $\hat{\phi}_2$ بواسطة $r(2)$ ، ونعيد هذه الطريقة بالنسبة للقيم المثلى لـ p . ثم، نسمي هذه السلسلة $(r(1),r(2),.....)$ بدالة الارتباط الذاتي الجزئية، وعلى العموم إذا كانت الدرجة الحقيقية للسيرورة هي p فإننا نلاحظ أن[1]:

$k>p$: $r(k)=0$ ، وبعبارة أخرى فإن دالة الارتباط الجزئية لنموذج AR(p) تنعدم بعد فجوة زمنية تساوي p.

3.5. النماذج المختلطة Mixed models ARMA(p.q):

1.3.5. نماذج ARMA(p.q) المستقرة:

هناك سيرورات عشوائية لا يمكن نمذجتها على أنها مجرد متوسط متحرك أو انحدار ذاتي فقط، بل يمكن أن تحتوي على خصائص النوعين من السيرورات معا. بحيث تشمل هذه النماذج على القسم الانحداري ذي الدرجة p وقسم المتوسطات المتحركة ذي الدرجة q، كما يظهر في الكتابة التالية:

1- Gouriéroux and Manfort (1983), p.149.

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q}$$

كما أن الشرط الضروري لاستقرار السيرورة ARMA(p.q) هو $\sum_{i=1}^{p} \phi_i < 1$ ، بحيث يكون متوسطها ثابتا عبر الزمن، ويُعطي النتيجة التالية: $\mu = \delta \Big/ \left(1 - \sum_{i=1}^{p} \phi_i\right)$.

ومن خصائص دالة الارتباط الذاتي للسيرورة ARMA(p.q) أنها تأخذ الشكل الانحداري بعد الفجوة الزمنية q أي تتناقص بشكل أسي انطلاقا من $k > q$.[1]

أما دالة الارتباط الجزئي فإنها تأخذ شكل دالة الارتباط الذاتي الجزئي لنموذج المتوسطات المتحركة بعد الفجوات الزمنية p، أي تتناقص بشكل أسي انطلاقا من $k > p$.[2]

ولنعتبر أبسط حالة وهي ARMA(1.1) على الشكل:

$$Y_t = \phi_1 Y_{t-1} + \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

وبوضع $\delta = 0$ تكون التباينات والتباينات المشتركة لهذه الأخيرة:

$$\gamma(0) = \text{var}(Y_t) = E[Y_t(\phi_1 Y_{t-1} + \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1})]$$

ومن أجل $|\phi_1| < 1$ ينتج في الأخير:

$$\gamma(0) = \sigma_\varepsilon^2 [1 + \theta_1^2 + 2\phi_1\theta_1] / (1 - \phi_1^2)$$

أما التباينات المشتركة فهي:

$$\gamma(1) = E(Y_t Y_{t-1}) = \phi_1 \gamma(0) - \theta_1 \sigma_\varepsilon^2$$
$$\gamma(2) = E(Y_t Y_{t-2}) = \phi_1 \gamma(1)$$
$$.......................................$$
$$\gamma(k) = E(Y_t Y_{t-k}) = \phi_1 \gamma(k-1) \quad : k \geq 2$$

من أجل السيرورة ARMA(p.q) لدينا:

1- Tenenhaus (1994), p. 295.
2- Bresson and Michaud (1995) , p.38.

$$\gamma(k) = E(Y_t Y_{t-k}) = \phi_1\gamma(k-1) + \phi_2\gamma(k-2) + ... + \phi_p\gamma(k-p) \quad : k \geq q+1$$
$$\rho(k) = \phi_1\gamma(k-1) + \phi_2\gamma(k-2) + ... + \phi_p\gamma(k-p) \quad : k \geq q+1$$

ونلاحظ أن q هو ذاكرة الجزء MA(q)، فمن أجل $k \geq q+1$ تأخذ دالة الارتباط الذاتي للسيرورة ARMA(p.q) خصائص السيرورة AR(p) فقط.

ندرس شروط الاستقرارية، لدينا الصيغة الرياضية للسيرورة ARMA(p.q):

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \delta + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + ... + \theta_q\varepsilon_{t-q}$$

و بإدخال معامل التأخير L فإن الصيغة تصبح:

$$(1 - \phi_1 L - \phi_2 L^2 - ...... - \phi_p L^p) y_t = (1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q)\varepsilon_t$$

أي:
$$\Phi(L) y_t = \theta(L)\varepsilon_t \quad \Leftrightarrow \quad ARMA(p,q)$$

حيث y_t هي انحراف Y_t عن وسطها، وإذا كانت Y_t مستقرة فإن $\Phi^{-1}(L)$ يجب أن تتقارب، ويتطلب ذلك أن تكون جذور المعادلة المُميزة تقع خارج دائرة الواحد Outside Unit Circle لتكون الحلول $L_1, L_2, ... L_p$ للمعادلة $\Phi(L) = 0$ كلها أكبر من الواحد (بالقيمة المطلقة)، وإذا تحقق ذلك نكتب المعادلة $\Phi(L) y_t = \theta(L)\varepsilon_t$ على الشكل:

$$y_t = \Phi^{-1}(L) \ .\theta(L)\varepsilon_t$$

ونقول عن y_t بأنها قابلة للقلب إذا استطعنا كتابة المعادلة على الشكل: $\theta^{-1}(L).\Phi(L) y_t = \varepsilon_t$. ومنه إذا استطعنا قلب السيرورة ARMA(p.q) إلى السيرورة AR(p) فقط، وإذا كانت Y_t قابلة للقلب، فإن $\theta^{-1}(L)$ يجب أن تتقارب بشرط أن تقع جذور المعادلة المميزة $\theta \ (L) = 0$ خارج دائرة الواحد.

- وكمثال نعتبر السيرورة MA(1)، والتي تكون معادلتها المميزة هي: $1 - \theta_1 L = 0$، ومنه فإن شرط وجود المقلوب:

$L = \frac{1}{|\theta_1|} > 0$ أو أن $|\theta_1| < 0$.

- أما من أجل MA(2) تكون المعادلة المميزة هي: $1-\theta_1 L-\theta_2 L^2=0$.

ومنه تكون جذورها على الشكل: $L=\frac{-\theta_1 \pm \sqrt{\theta_1^2+4\theta_2}}{2\theta_2}$.

إن القيمتين L_1 و L_2 يجب أن تقعا خارج دائرة الواحد، والتي تستلزم أن:

$\theta_2+\theta_1<1, -\theta_2-\theta_1<1, |\theta_2|<1$.

2.3.5. نماذج ARMA(p.q) غير المستقرة ARIMA(p,d,q) models:

إذا كانت السلسلة الزمنية الأصلية غير مستقرة فيقال عليها أنها متكاملة Integrated or Nonstationary، وإذا تعين الحصول على فروقات السلسلة d مرة حتى تصبح مستقرة، يقال عندئذ أن السلسلة الأصلية متكاملة من الدرجة d $I(d)$.

وبعبارة أخرى نقول أن Y_t هي سلسلة متجانسة و غير مستقرة متكاملة من الدرجة d إذا وجدت $W_t=\nabla^d Y_t$ سلسلة مستقرة جديدة. ومنه يمكن أن نُنَمذِّج السلسلة الجديدة W_t كأنها سيرورة ARMA(p,q)، في هذه الحالة ينتج أن Y_t هي سيرورة ARIMA(p,d,q)[1]، ونسمي ذلك بنموذج الانحدار الذاتي والمتوسط المتحرك المُتكامل، هذا الأخير بالإضافة إلى الدرجتين p و q فإنه يتميز بدرجة ثالثة d.. يكتب على الشكل:

$$\Phi(L)(1-L)^d Y_t=\delta+\theta(L)\varepsilon_t$$

أو:

$$\Phi(L)\nabla^d Y_t=\delta+\theta(L)\varepsilon_t$$

ويلاحظ أن وسط $W_t=(1-L)^d Y_t$ المستقر هو $\mu_w=\delta\Big/\left(1-\sum_{i=1}^{p}\phi_i\right)$، وبالتالي إذا كانت $\delta=0$ فإن السلسلة المُتكاملة W_t سوف يكون لها اتجاه عام محدّد البناء. وكمثال فإذا كان النموذج ARIMA(1,1,1) فهذا يعني أنه يتعين الحصول على الفروقات الأولى

1- ARIMA هي اختصار لـ: Autoregressive Integrated Moving Average Process

للسلسلة الأصلية ثم نجري عليها بعد ذلك تقدير ARMA، ذلك لأن هذا الأخير لا يُجرى إلا على سلسلة مستقرة، وتكون صيغة النموذج عندئذ:

$$.\ \Delta Y_t = \phi_1 \Delta Y_{t-1} + \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

وعُموما يمكن القول أن: $ARIMA(p,0,q) = ARMA(p,q)$ وتكون السلسلة الأصلية مستقرة، وأيضا $ARIMA(p,0,0) = AR(p)$ و $ARIMA(0,0,q) = MA(q)$.

3.3.5. النمـاذج الموسميـة المختلطـة SARIMA " Seasonal Autoregressive Integrated Moving Average":

تتميز السلاسل الزمنية في الواقع بوجود المركبة الموسمية، الشيء الذي يؤدي إلى ارتفاع كل من p و q، وبالتالي تصعب عملية تقديرها، ولأجل ذلك وُضِع نموذج يسمى بالنموذج المختلط ذي المركبة الموسمية SARIMA(p,d,q). ويمكن التعبير عنه رياضيا كما يلي:

$$\phi(L)\Phi(L^s)\nabla^d \nabla_s^D Y_t = \theta(L)\Theta(L^s)\varepsilon_t$$

حيث:

$$\Phi(L^S) = 1 - \phi_1 L^S - \phi_2 L^{2S} - \ldots\ldots\ldots\ldots - \phi_p L^{pS}$$
$$\Theta(L^S) = 1 - \theta_1 L^S - \theta_2 L^{2S} - \ldots\ldots\ldots\ldots - \theta_q L^{qS}$$
:

يُمثل $\nabla_s^D = (1-L^s)^D$ الفروقات الموسمية من الدرجة D و $\nabla^d = (1-L)^d$ الفروقات المتتالية من الدرجة d اللذان يستخدمان لتحقيق استقرارية Y_t.

6. منهجية بوكس– جينكتر في بناء نماذج السلاسل الزمنية الخطية:

اهتم Box and Jenkins (1976) بجمع بعض التقنيات المستعملة في السلاسل الزمنية للمساعدة على تحديد درجة النموذج وتقدير معالمه، ثم اقتراح بعض الطرق للتأكد من صلاحية النموذج لأخذ شكله النهائي.

رأينا أن السلسلة الزمنية غير المستقرة والمتجانسة يمكن أن تُنَمذَّج على شكل ARIMA(p,d,q)، فيكمن المشكل التطبيقي إذن في كيفية اختيار القيم الثلاثة p,d,q أي لتحديد شكل هذا النوع من النماذج، نختبر دالتي الارتباط الذاتي البسيط والجزئي للسلسلة والمشكل فيها هو تحديد درجة التكامل d من أجل الحصول على السلسلة المستقرة، ومنه لتحديد القيمة العددية المناسبة لـ d. نستعمل الفكرة القائلة بأن الارتباط الذاتي $\rho(k)$ بالنسبة للسلاسل الزمنية المستقرة، يجب أن يقترب تدريجيا من الصفر كلما كبر عدد التباطؤات k، ولمعرفة ذلك نعتبر نموذج السيرورة ARMA(p.q)، حيث أن دالة الارتباط للجزء MA(q) تؤول الى الصفر عندما تكون $k > q$، لأن لهذا النمط ذاكرة تساوي q فترة فقط، ومنه إذا كانت Y_t تتبع السيرورة MA(q)، فإن $\rho(k) = 0$ من أجل $k > q$ وتتناقص دالة الارتباط الذاتي للجزء AR(p) من السيرورة ARMA(p.q) المستقرة هندسيا.

إن طريقة تحديد قيمة d هي مباشرة، حيث ننظر إلى دالة الارتباط الذاتي للسلسلة الأصلية أو نختبر الجذر الوحدوي ونحدد ما إذا كانت مستقرة أم لا، فإذا حدث وإن كانت غير مستقرة نلجأ إلى استعمال تقنية الفروقات على السلسلة لكي نجعلها مستقرة. ونعيد هذه الطريقة حتى نصل إلى القيمة d التي تجعل السلسلة مستقرة أي $W_t = (1-L)^d Y_t$ وهذا معناه أن دالة الارتباط الذاتي $\rho(k)$ تؤول إلى الصفر لما تكون k كبيرة. وفي هذه الحالة نقول أن Y_t قابلة للمكاملة من الدرجة d ويكون بذلك الفرق الـ $(d-1)$ غير مستقر[1].

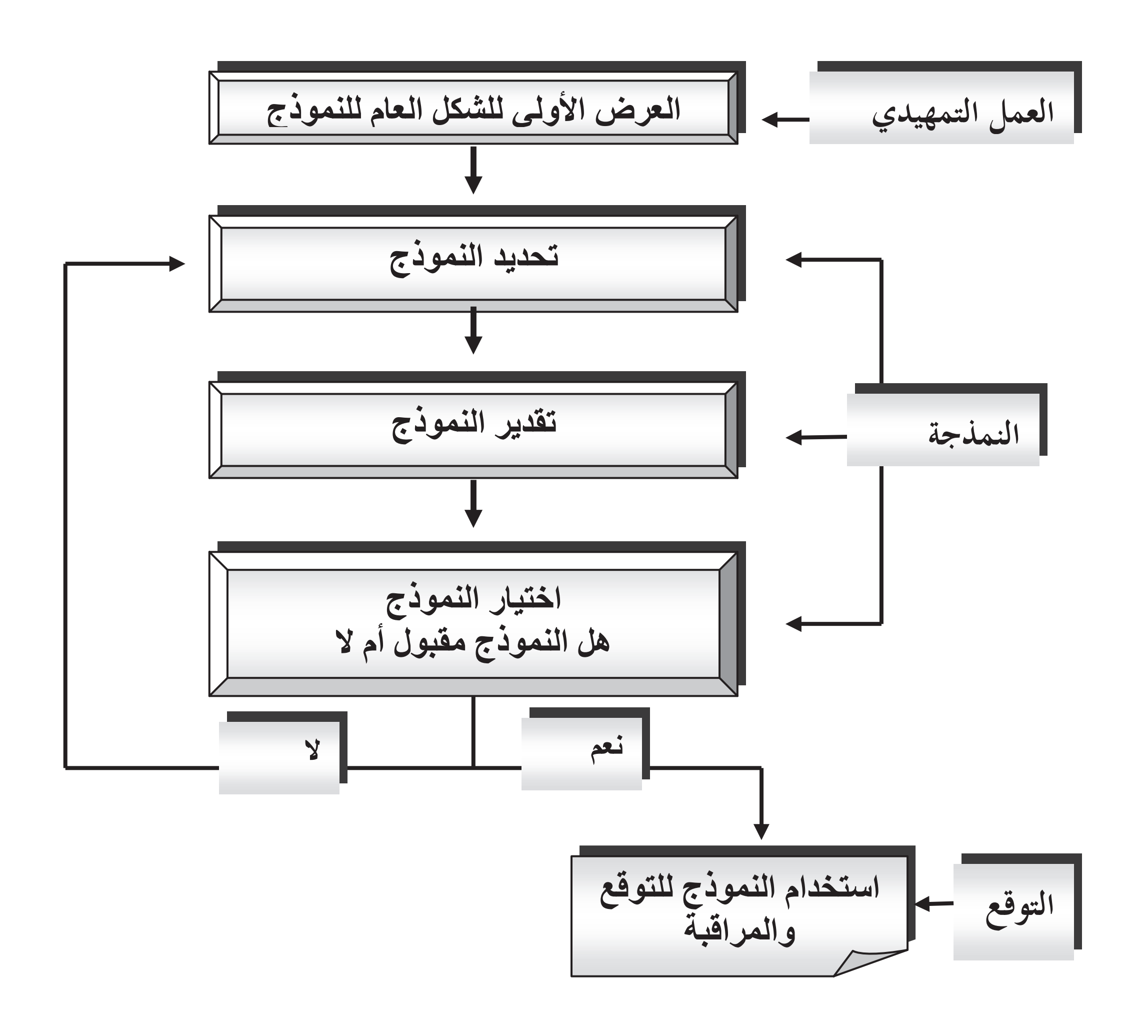
بعد تحديد قيمة d، يمكن استعمال السلسلة المستقرة $W_t = (1-L)^d Y_t$ لاختبار كل من دالة الارتباط الذاتي البسيط و الجزئي لتحديد p و q، وإذا كان لكل من الجزء AR و MA درجات عليا، يمكن استعمال الطريقة التجريبية لكل من p و q، ثم نتأكد من ذلك التجريب بعد تقدير معالم النموذج ARMA(p,.q) للسلسلة المُحوَّلة.

1- بوشة محمد، 2000، ص 90.

ويرى (1976) Box and Jenkins أن النماذج الديناميكية الخطية المقدرة والتحليلات النظرية المرافقة لها لا تعطينا شكل النموذج فقط، وإنما نحصل أيضا على المعالم المقدرة.

يبين المخطط التالي الخطوات العملية حسب منهجية بوكس وجينكز لبناء نموذج خطي لسلسلة زمنية واحدة، بغرض التوقع والمراقبة في المدى القصير:

الشكل رقم (10)[1] : منهجية بوكس– جينكز في بناء نماذج السلاسل الزمنية الخطية



1- Box and Jenkins (1976), p. 19

من خلال هذا المخطط يتبن أنه هناك أربعة خطوات يتعين إتباعها حـ ـتى نسـ ـتخدم منهجية بوكس-جينكز في التنبؤ، تتمثل فيما يلي:

1. مرحلة التعرف (التحديد) Identification
2. مرحلة التقدير Estimation
3. مرحلة الفحص (المراقبة والضبط) التشخيصي Diagnostic
4. مرحلة التنبؤ Prediction

1.6. مرحلة التعرف (التمييز):

إن أصعب مرحلة في بناء نماذج السلاسل الزمنية الخطية هي مرحلة التمييز، حيث يمكن الحصول على عدة بدائل للنماذج الممكنة، كما يمكن رفض النموذج الأولي المختار في مرحلة الفحص والاختبار[1]. إذا أظهرت السلسلة Y_t اتجاها عاما قويا فإن حساب الفروقات من الدرجة الأولى أو الثانية سوف يؤدي الى استقرار السلسلة غالبا W_t، ولتحديد درجة الانحدار الذاتي p، ودرجة المتوسط المتحرك q نستخدم دالتي الارتباط الذاتي والجزئي.

إذا كان شكل الارتباط يقع داخل حدود فترة الثقة 95 % منذ البداية، فإن معامل الارتباط الذاتي (ACF) لا يختلف جوهريا عن الصفر فهذا يعني أن السلسلة مستقرة ومتكاملة من الدرجة 0، في هذه الحالة نجري تحليلاتنا على القيم الأصلية للمتغير Y_t، دون إجراء تحويلات عليها، أما إذا اتضح أن شكل الارتباط الذاتي يقع خارج مجال الثقة 95 % في فترة طويلة ومعاملات الارتباط الذاتي تختلف معنويا عن الصفر من أجل k كبير نسبيا، فإن السلسلة Y_t تكون غير مستقرة، في هذه الحالة يجب إجراء الفروقات من الدرجة الأولى ثم نجري عليها نفس التحليل مرة أخرى حتى نصل إلى سلسلة مستقرة.

1- تومي صالح، ج(2)، ص 183.

بعد الحصول على الاستقرار فإنه يمكن دراسة الارتباطات الذاتية والارتباطات الذاتية الجزئية للسلسلة المستقرة لتساعدنا على تمييز نوعية السلوك الخاص بالانحدار الذاتي أو المتوسط المتحرك أو لكليهما معا، ولاختيار النموذج نقترح المعايير التالية:

1.1.6. معيار Hannan-Rissanen:

حسب Hannan and Rissanen (1982)، إذا كانت لدينا T مشاهدة (مع T كبيرة بدرجة كافية) وتوصلنا إلى درجة معقولة من الفروقات للسيرورة، فإن السلسلة المحولة W_t ذات متوسط معدوم. لدينا نموذج ARMA(p,q):

$$\Phi(L)W_t = \theta(L)\varepsilon_t$$

- نحاول أولا تفريقها بواسطة الانحدار الذاتي من الدرجة s المطلوب تحديدها وتأخذ الشكل:

$$W_t = \phi_{s1}W_{t-1} + \phi_{s2}W_{t-2} + \dots\dots\dots\dots + \phi_{ss}W_{t-s} + \varepsilon_t$$

- بوجود الارتباطات r ، فإن المعالم ϕ_{sj} يمكن تقديرها، بالتراجع وفقا لطريقة Durbin (1960) ، والتي تعطي:

$$\hat{\phi}_{11} = r_1 \quad , \hat{\phi}_{ss} = \frac{r_s - \sum_{j=1}^{s-1} \phi_{s-1,j}\, r_{s-j}}{1 - \sum_{j=1}^{s-1} \phi_{s-1,j}\, r_j}$$

$$\hat{\phi}_{sj} = \hat{\phi}_{s-1,j} - \hat{\phi}_{s-1,s-j} \qquad : j = 1,2,\dots\dots,s-1$$

حيث $\hat{\phi}_{ss}$ هي الارتباطات الذاتية الجزئية.

2.1.6. معيار Akaike (تحديد الدرجة المقربة للانحدار الذاتي):

يكون تحديد القيمة المناسبة لـ s (الدرجة المقربة للانحدار الذاتي) عن طريق استعمال معيار Akaike (1969) أي نختار قيمة s عندما يكون هذا المعيار أصغر ما يمكن:

$$AIC = TLog\,\hat{\sigma}_s^2 + 2s$$

حيث أن AIC هو معيار المعلومات لـ . Akaike، وs هو عدد المعالم، أما إذا استعملنا عدة عينات مختلفة الحجم بالنسبة لنفس السلسلة Y_t أو W_t، فإننا نستعمل معيار المعلومات المرجح والذي يعطي أصغر قيمة للمقدار:

$$NAIC = .Log\hat{\sigma}_s^2 + 2s/T$$

حيث أن $\hat{\sigma}_s^2$ هو مقدار تباينات الأخطاء لنماذج الانحدارات الذاتية المقدرة والتي يمكن إيجادها بالتراجع من:

$$\hat{\sigma}_1^2 = \left(1 - r_1^2\right)\sum_{t=1}^{T}\frac{W_t^2}{T} \ , \quad \hat{\sigma}_s^2 = \left(1 - \hat{\phi}_{ss}^2\right)\hat{\sigma}_{s-1}^2$$

إن الهدف من تقدير الانحدار الذاتي المقرب هو الحصول على مقدرات للتذبذبات Innovations ε_t، حيث إذا كانت القيمة المختارة لـ . s هي s^*، يمكن أن نستعمل لذلك البواقي على الشكل:

$$\hat{\varepsilon}_t = W_t - \hat{\phi}_{s^*1}W_{t-1} - \hat{\phi}_{s^*2}W_{t-2} - .... + \hat{\phi}_{s^*s^*}W_{t-s^*} + \varepsilon_t$$

- ويمكن استعمال هذه البواقي مكان التذبذبات المؤخرة ε_{t-1} في التشكيلة ARMA(p,q)، ومنه يمكن أن نكتب:

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + ... + \phi_p W_{t-p} + \varepsilon_t + \theta_1\hat{\varepsilon}_{t-1} + \theta_2\hat{\varepsilon}_{t-2} + ... + \theta_q\hat{\varepsilon}_{t-q}$$

إن مزايا هذه المعادلة هو أنه يمكن تقدير المعالم ϕ_j, θ_i $\left(i = 1....p \ , \ j = 1....q\right)$ نسبيا بواسطة المربعات الصغرى العادية من أجل القيم p و q حيث أن Hannan and Rissanen (1982) يقترحان اختيار القيم الخاصة بـ . p و q التي تحقق أصغر قيمة للعبارة التالية[1]:

$$HR = Log\sigma_{p,q}^2 + \frac{(p+q)LogT}{T}$$

1- Hannan.E.J and Rissannen (1982), p. 81.

ولقد أثبت جدية هذه الطريقة عدة باحثين عبر تجارب مُطبقة على مختلف العينات المطورة، ويقترح Hannan and Kavlier (1984) تحويلات مختلفة للطريقة الأصلية والتي تعطي مقدرات متسقة.

وقبل التطرق إلى موضوع التقدير، نود تلخيص مجمل الخطوات الضرورية أثناء العمل التطبيقي المتمثل في المراحل التالية[1]:

1. تكون دالة الارتباط الذاتي (AC) مؤشرا مهما لكشف عدم استقرارية سلسلة زمنية، وهذا عندما لا تنعدم هذه الدالة بعد عينة تعادل $\frac{T}{4}$ (ربع عدد المشاهدات) نظريا، بينما تطبيقيا يجب أن تقع معاملات هذه الدالة داخل مجال ثقة مناسب حتى تكون السلسلة مستقرة (وإلا فلا)، وهنا نكون بصدد دراسة النماذج المركبة، كما أنها تعتبر كاشفا مهما للمركبة الموسمية من خلال القمم والنتوءات التي تظهر في شكل منتظم على هذه الدالة.
2. بالنسبة لنماذج المتوسطات المتحركة من الدرجة q تنعدم معاملات الارتباط الذاتي معنويا مباشرة بعد الدرجة q، بينما دالة الارتباط الجزئية تبقى متدهورة أي متناقصة بعد هذه الفترة ولكنها لا تنعدم: $\forall k > q : \hat{\rho}(k) = 0$.
3. بالنسبة لنماذج الانحدار الذاتي من الدرجة p، فإن معاملات الارتباط الذاتي الجزئية تنعدم معنويا مباشرة بعد هذه الدرجة، بينما تبقى دالة الارتباط الذاتي متناقصة ولكنها لا تنعدم بنفس السرعة: $\forall k > p : \hat{r}(k) = 0$
4. أما النماذج المختلطة فإن الدالتين تبقيان مستمرتين في التدهور ولكنهما لا تنعدمان معنويا عند الدرجتين المذكورتين سابقا. في الحالة نستخدم المعايير التي ينبغي أن تكون أصغر ما يمكن لتحديد الدرجتين p و q مثل معيار AIC أو BIC....الخ.

1- مولود حشمان، ص 145.

والجدول التالي يلخص الحالات الثلاثة الأخيرة:

الجدول (3): تطور طبيعة النموذج وفق منحنى الارتباط الذاتي

نوع النموذج	ACF	PACF
MA(q)	تنعدم معنويا بعد الفترة q	غير منعدمة معنويا DIES OUT
AR(p)	غير منعدمة معنويا DIES OUT	تنعدم معنويا بعد الفترة p
ARMA(p,q)	غير منعدمة معنويا DIES OUT	غير منعدمة معنويا DIES OUT

مثال 4:

بالنظر إلى الشكل (8) الذي يمثل منحنيات دوال الارتباط البسيطة والجزئية للسلسلة المستقرة، نلاحظ أن معامل الارتباط $\rho(1)$ يختلف معنويا عن الصفر (أي يقع خارج مجال الثقة) ومن أجل $k > 1$ كل معاملات الارتباط الذاتي تنعدم معنويا، وهي الحالة التي توافق نموذج MA(1)، كما نلاحظ أيضا أن معامل الارتباط الجزئي $r(1)$ يختلف معنويا عن الصفر ومن أجل $k > 1$ كل معاملات الارتباط الجزئي تنعدم معنويا، وهي الحالة التي توافق نموذج AR(1).

وفقا لهذه النقاط تكون الصيغة الرياضية المثلى للنموذجين المرشحين المعرفين للسلسـ ـلة المستقرة من الشكل:

$$ARIMA(0,1,1): \nabla Y_t = \delta + (1+\theta_1 L)\varepsilon_t$$
$$ARIMA(1,1,0): (1-\phi_1 L)\nabla Y_t = \delta + \varepsilon_t$$

وبعد تقدير هذين النموذجين، يكون النموذج المختار هو الذي يُعطي أحسن توفيقة بين المعايير Akaike، Schwarz، أي تصغير هذين لمعيارين.

2.6. مرحلة تقدير معالم النموذج:

بعد الانتهاء من مرحلة التعرف على النموذج بتحديد الرتب (أو الدرجات) p، d و q. يمكننا الانتقال إلى المرحلة التقنية الموالية والمتمثلة في مرحلة التقدير لمعالم النموذج.

1.2.6. تقدير معالم نموذج الانحدار الذاتي AR:

في هذا النوع من النماذج، وبعد تحديد الدرجة p، يصبح من الميسور تقدير معالمه $(\phi_p,......., \phi_2, \phi_1)$، وذلك باستعمال إحدى الطرق التالية:

أ. **طريقة معادلات يول– ولكر Yule-Walker :**

ترتكز هذه الطريقة على معادلات يول–ولكر التي تحدثنا عنها سابقا من خلال معاملات الارتباط الذاتي لتقدير معالم النموذج، حيث أن المقدرات في حالة نماذج AR(p) تكون فعالة. لدينا:

ففي حالة AR(p) مثلا تكون لدينا p معادلة ليول وولكر:

$$\rho(1) = \phi_1 + \phi_2\rho(1) + ... + \phi_p\rho(p-1)$$

$$\rho(1) = \phi_1\rho(1) + \phi_2 + ... + \phi_p\rho(p-2)$$

....

$$\rho(p) = \phi_1\rho(p-1) + \phi_2\rho(p-2) + ... + \phi_p$$

تكتب هذه المعادلات على الشكل المصفوفي:

$$\begin{pmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{pmatrix} = \begin{pmatrix} 1 & \rho(1) & \cdots & \rho(p-1) \\ \rho(1) & 1 & \cdots & \rho(p-2) \\ \vdots & \vdots & 1 & \vdots \\ \rho(p-1) & \rho(p-2) & \cdots & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix}$$

وبتعويض المعالم بمقدراتها، نحصل على الشكل المختصر:

$$R = A \times \hat{\Phi}$$

ومنه:

$$\hat{\Phi} = A^{-1} \times R$$

ب. الطريقة الانحدارية:

ليكن نموذج AR(p):

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t$$

وبكتابتها على الشكل المصفوفي:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & Y_1 & 0 & \cdots & \vdots \\ 1 & \vdots & \vdots & \cdots & \vdots \\ 1 & Y_{T-1} & Y_{T-2} & \cdots & Y_{T-p} \end{pmatrix} \times \begin{pmatrix} \phi_0 \\ \phi_1 \\ \vdots \\ \phi_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{pmatrix}$$

فنحصل على الكتابة المختصرة:

$$Y = X \times \Phi + \varepsilon$$

حيث $X(T, p+1)$ مصفوفة المتغيرات المستقلة، $Y(T,1)$ المتغير التابع، $\Phi(p+1,1)$ شعاع المعالم الواجب تقديرها، $\varepsilon(T,1)$ شعاع الأخطاء. نذكر فقط أننا سنفقد p مشاهدة، فقمنا بتعويض تلك القيم المفقودة بـ 0 . وتحت فرضيات معينة معروفة يمكن تقدير شعاع المقدرات بطريقة المربعات الصغرى العادية كما يلي:

$$\hat{\Phi} = \left(X'X\right)^{-1} X'Y$$

2.2.6. تقدير معالم المتوسطات المتحركة والمختلطة:

تعتبر هذه النماذج MA(q) و ARMA(p,q) أعقد بكثير من حيث التقدير من النماذج الانحدارية، كونها غير خطية في المعالم من جهة، وعدم مشاهدة متغير الأخطاء من جهة أخرى.

فهدف التقدير هنا هو تحديد معالم القسم الانحداري وقسم المتوسطات المتحركة ARMA(p,q) معا، أو معالم قسم المتوسطات المتحركة لوحدها في نموذج MA(q) ففي حالة النموذج المختلط التالي:

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \dots\dots\dots - \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots\dots\dots + \theta_q \varepsilon_{t-q}$$

أي:
$$\Phi(L)Y_t = \theta(L)\varepsilon_t$$

حيث $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots\dots - \phi_p L^p$ و $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots\dots + \theta_q L^q$

بافتراض إمكانية قلب المعامل $\theta(L)$ فإن:
$$\varepsilon_t = \theta^{-1}(L)\,\Phi(L)Y_t$$

إذًا فإن أي طريقة تقدير، يجب أن تأخذ بعين الاعتبار فكرة تصغير (أو تدنئة) مجموع مربعات البواقي، أي:

$$Min \sum \hat{\varepsilon}_t^2 = s(\hat{\phi}, \hat{\theta})$$

حيث:
$$\hat{\varepsilon}_t = \hat{\theta}^{-1}(L)\hat{\Phi}(L)Y_t$$

لقد رأينا إمكانية وسهولة تقدير معالم هذه العلاقة في حالة غياب الطرف MA(q)، بينما في حالة وجودها لوحدها أو مع مركبة الانحدار الذاتي AR(p)، فإن هذه العلاقة تصبح غير خطية المعالم، وبالتالي تتطلب طريقة تقدير تكرارية Non Linear Iterative Routine، ومن بين هذه الطرق:

أ. **طريقة البحث التشابكي Grid-Search**:

لتوضيحها ندرج النموذج المختلط البسيط التالي ARMA(1,1):

$$Y_t - \phi_1 Y_{t-1} = \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

إذن:
$$(1 - \phi_1 L)Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

ومنه:
$$Y_t = \frac{1}{(1 - \phi_1 L)}(1 + \theta_1 L)\varepsilon_t$$

نضع:
$$v_t = \frac{1}{(1 - \phi_1 L)}\varepsilon_t$$

يصبح:
$$v_t = \phi_1\, v_{t-1} + \varepsilon_t$$

نلاحظ عند هذه العلاقة الأخيرة، أنه لو توفرت قيم الشعاع v_t، فإننا نستطيع تقـ ـدير المعلمة ϕ بطريقة المربعات الصغرى العادية، ولكن بسبب عدم مشاهدتها نلجأ إلى العملية التالية حيث نستطيع كتابة:

$$Y_t = \frac{1}{(1-\phi_1 L)}\varepsilon_t + \frac{\theta_1}{(1-\phi_1 L)}\varepsilon_{t-1}$$

أي أن:

$$Y_t = v_t + \theta_1 \, v_{t-1}$$

ومن هذه المعادلة وبتعويض θ_1 بقيمها، والتي تقع ضمن المجال $|\theta_1| < 1$ من أجل شرط إمكانية قلب النموذج، وبتوفير القيم البدائية لـ v_t. أو جعلها مساوية للصفر، (في هذا المثال $v_0 = 0$) نحصل على: $v_t = Y_t + \theta_1 \, v_{t-1}$ ، إذن تبدأ العملية بالتكرار الأول 1^{st} iteration وذلك باختيار مثلا $\theta_1 = -0.9$ ونسميها $\theta_1^{(1)}$ وكما يلي[1]:

$$\begin{aligned}
t &= 1 \quad : v_1^{(1)} = Y_1 \\
t &= 2 \quad : v_2^{(1)} = Y_2 - \theta_1^{(1)} v_1^{(1)} \\
t &= 3 \quad : v_3^{(1)} = Y_3 - \theta_1^{(1)} v_2^{(1)} \\
&\vdots \\
t &= T \quad : v_n^{(1)} = Y_n - \theta_1^{(1)} v_{n-1}^{(1)}
\end{aligned}$$

حيث: $V'^{(1)} = \left[v_1^{(1)}, v_2^{(1)}, ..., v_n^{(1)}\right]$ وبتعويض هذا الشعاع الناتج، نستطيع تقدير المعلمة ϕ_1 باستعمال طريقة OLS:

$$\hat{\phi}_1^{(1)} = \frac{\sum_t v_t^{(1)} v_{t-1}^{(1)}}{\sum_t \left[v_{t-1}^{(1)}\right]^2}$$

ثم نقوم بحساب مجموع مربعات البواقي المقابلة للمعلمين $\left(\theta_1^{(1)}, \hat{\phi}_1^{(1)}\right)$ كالآتي:

$$\sum_t \hat{\varepsilon}_t^2 = \sum_t \left[v_t^{(1)} - \hat{\phi}^{(1)} v_{t-1}^{(1)}\right]^2$$

1- الرقم الذي بين قوسين يمثل دليل التكرار.

ونسمي مجموع المربعات هذه بالرمز المتعارف عليه والموافق للتكرار الأول $RSS^{(1)}$ و نعيد العملية للمرة الثانية (التكرار الثاني) وفق المراحل السابقة والتي نختصرها فيما يلي:

- حساب الشعاع $V^{(2)}$ باستعمال قيمة θ_1 الموالية $\theta_1 = -0.8$ مثلا (إذا كان مقدار الخطوة يعادل 0.1) $\hat{\phi}_1^{(2)} = \frac{\sum_t v_t^{(2)} v_{t-1}^{(21)}}{\sum_t \left[v_{t-1}^{(2)}\right]^2}$

- تقدير المعلمة:

- حساب مجموع مربعات البواقي: $RSS^{(2)} = \sum_t \hat{\varepsilon}_t^2 = \sum_t \left[v_t^{(2)} - \hat{\phi}^{(1)} v_{t-1}^{(2)}\right]^2$

ونعيد هذه العملية حتى نغطي كاملا مجال التعويض لـ θ_1. وحتى نتحصل على المعالم التي تدني RSS.

نشير هنا إلى أن هذه الطريقة تصبح غير مرغوب فيها لما يتجاوز عدد معالم قسم المتوسطات المتحركة درجتين $q > 2$، نظرا لصعوبة عملية الحساب من جهة وكذا عدم اتساق المعالم في هذه الحالة.

ب. طريقة غوس– نيوتن Gauss- Newton :

تعتمد هذه الطريقة كذلك على تدنيه أو تصغير مجموع مربعات البواقي، حيث:

$$\hat{\varepsilon}_t = \theta^{-1}(L)\Phi(L)Y_t$$

وبما أن هذه المعادلة غير خطية المعالم، فإنه لا يمكن تقديرها بواسطة التطبيق المباشر للمربعات الصغرى العادية، للحصول على $\hat{\phi},\hat{\theta}$ يمكن استعمال طريقة التقدير غير الخطي لـ Gauss-Newton ، مستعملين نشر تايلور Taylor لضبط المعادلة السابقة في شكل خطي، حول قيمة انطلاق معينة للشعاعين ϕوθ، نعيد هذه السيرورة حتى يحدث التقارب. فإذا أخذنا نموذج السيرورة ARMA(1,1): $Y_t = \phi_1 Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$ مع ε_t مستقلة ومتماثلة التوزيع مهما تكن t، ومن أجل $|\phi_1| < 1$ نضرب طرفي المعادلة بالمقدار $\theta^{-1}(L)$ فنجد: $\theta^{-1}(L)Y_t = \theta^{-1}(L)\phi_1 Y_{t-1} + u_t$

إن المشكل الأساسي في هذه المعادلة هو كيفية شرح المتغير المحول $\theta^{-1}(L)Y_t$ الذي هو عبارة عن مجموع الترجيحات للقيم الحالية والماضية للسلسلة Y المحتوية على قيم العينة السابقة والتي تكون غير ملاحظة، وإذا فرضنا أن كل قيم العينة السابقة للسلسلة Y مساوية للصفر، تصبح العملية بسيطة، فانطلاقا من هذه الفرضية، تكون السلسلة المحولة هي:

$$Y_t^* = \theta^{-1}(L)Y_t \qquad : t = 1,2,.....,T$$

والتي هي على الشكل:

$$\begin{aligned} Y_1^* &= Y_1 \\ Y_2^* &= Y_2 + \theta_1 Y_1 \\ Y_3^* &= Y_3 + \theta_1 Y_2 + \theta_1^2 Y_1 \\ &\vdots \\ Y_n^* &= Y_n + \theta_1 Y_{n-1} + \theta_1^2 Y_{n-2} + ... + \theta_1^{t-1} Y_1 \end{aligned}$$

وبالتراجع نستنتج أن السلسلة على الشكل: $Y_t^* = Y_t + \theta_1 Y_{t-1}^* \quad : t = 1,2....,T$ مع $Y_0^* = 0$.

يمكن إعادة كتابة المعادلة $\theta^{-1}(L)Y_t = \theta^{-1}(L).\phi_1 Y_{t-1} + u_t$ على الشكل: $Y_t^* = \phi_1 Y_{t-1}^* + \varepsilon_t \quad : t = 1,2....,T$، حيث أن هذه المعادلة الأخيرة خطية في ϕ_1، وإذا كانت θ_1 معطاة فإن قيم السلسلة المحولة Y_t^* تصبح معروفة ليكون التطبيق المباشر لقانون المربعات الصغرى العادية يعطي مقدرا متسقا لـ ϕ_1. عمليا، تكون θ_1 غير معروفة، ومنه نضطر لتطبيق التقدير غير الخطي، وأبسط طريقة للحصول على المقدرات غير الخطية $\hat{\theta}_1$ و $\hat{\phi}_1$ هي استعمال طريقة البحث بمجال، حيث أن تطبيق قانون المربعات الصغرى، من أجل قيمة θ_1 في مجال مختار ويحقق الشرط $|\phi_1| < 1$، يعطي مقدرات متسقة، فمثلا نختار المجال $\theta \in]-1, +1[$ ونغير تدريجيا قيمة θ_1 في هذا المجال، مطبقين في كل مرة، قانون المربعات الصغرى من أجل الحصول على $\hat{\phi}_1$، ثم نختار قيمة θ_1 التي تحقق أصغر قيمة لمجموع البواقي للانحدار.

لكن هذه العملية تصبح مملة وتأخذ وقتا أكبر لما نواجه سيرورات ذات درجة $q > 2$، لذا يفضل أغلب الإحصائيين استعمال طرق التدنئة مثل طريقة Gauss- Newton للمربعات الصغرى غير الخطية.

حيث من خلال المعادلات السابقة نستطيع كتابة:

$$\varepsilon_t = \theta^{-1}(L)\Phi(L)Y_t = Y_t^* - \phi_1 Y_{t-1}^* \quad : t = 1,2....,T \qquad ....(7)$$

حيث أن: $\Phi(L)Y_t = Y_t - \phi_1 Y_{t-1}$

وما دام ε_t في (7) ليس خطيا في ϕ_1 و θ_1، فنستعمل نشر تايلور للسلسلة ε_t حول البواقي المقدرة $\hat{\varepsilon}_t$، لنجد:

$$\varepsilon_t = \hat{\varepsilon}_t + [\partial\varepsilon_t / \partial\phi_1](\phi_1 - \hat{\phi}_1) + [\partial\varepsilon_t / \partial\phi_1](\theta_1 - \hat{\theta}_1) + R_1 \qquad ....(8)$$

وبوضع $R_1 = 0$ نجد: $\hat{\varepsilon}_t = \frac{-\partial\varepsilon_t}{\partial\phi_1}(\phi_1 - \hat{\phi}_1) - \frac{\partial\varepsilon_t}{\partial\theta_1}(\theta_1 - \hat{\theta}_1) + \varepsilon_t$

يمكن اعتبار هذه الصيغة على أنها انحدار خطي حيث أن $\hat{\varepsilon}_t$ متغير تابع والمشتقات الجزئية $\frac{-\partial\varepsilon_t}{\partial\phi_1}$ و $\frac{-\partial\varepsilon_t}{\partial\theta_1}$ بجوار $\hat{\phi}_1$ و $\hat{\theta}_1$ على الترتيب، هي المتغيرات المفسرة (المستقلة)، إن الانحدار الناتج سوف يقدر القيم المراجعة للحصول على مقدرات جديدة للمعلمتين ϕ_1 و θ_1 ويحدث ذلك عن طريق تقسيم المشتقات في المعادلة (8) مستعملين المعادلة (7) في كل خطوة مراجعة والتي تعطي: $\frac{\partial\varepsilon_t}{\partial\phi_1} = -Y_{t-1}^*$. أما بالنسبة لـ $\frac{\partial\varepsilon_t}{\partial\theta_1}$ فنكتب: $Y_t = \phi_1 Y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1}$ لنجد أن: $\varepsilon_t = Y_t - \phi_1 Y_{t-1} + \theta_1\varepsilon_{t-1}$ وما دام Y_t و Y_{t-1} ملاحظتين من أجل كل t، فإن: $\frac{\partial\varepsilon_t}{\partial\theta_1} = \varepsilon_{t-1} + \theta_1\frac{\partial\varepsilon_{t-1}}{\partial\theta_1}$ ونستعمل هذه الصيغة بالتراجع مع اعتبار $\frac{\partial\varepsilon_t}{\partial\theta_1} = 0$ لما $t < 1$، ويمكن إعادة كتابتها على الشكل: $(1 - \theta L)\frac{\partial\varepsilon_t}{\partial\theta_1} = \varepsilon_{t-1}$.

نعرف المتغير المحول على الشكل: $\varepsilon_{t-1}^* = \frac{-\partial\varepsilon_t}{\partial\theta_1} = \frac{-\partial\varepsilon_{t-1}}{(1 - \theta L)}$.

ومن تم تتطلب طريقة Gauss-Newton تحديد البواقي المقدرة ε_t^* في المتغيرات Y_{t-1}^* و $-\varepsilon_{t-1}^*$، لتقدير التراجعات للمقدرات $\hat{\phi}_1$ و $\hat{\theta}_1$ على الترتيب. إن المتغيرات المكونة أعلاه يجب مراجعتها عند كل مرحلة من سيرورة التكرار، لأنها تعتمد على المقدرات الحالية، ونواصل العملية حتى تقترب المقدرات من الصفر.

3.6. مرحلة الاختبار Diagnostic Checking:

بعد الانتهاء من مرحلتي تحديد وتقدير النموذج، نود التطرق إلى المرحلة الثالثة من عملية النمذجة، وهي اختبار قوة النموذج الإحصائية ثم التنبؤية في مرحلة لاحقة، وهذه المرحلة تتطلب الخطوات التالية:

1.3.6. اختبار دالة الارتباط الذاتي للسلسلة:

نقارن دالة الارتباط الذاتي للسلسلة الأصلية مع تلك الخاصة بالسلسلة المقدرة، فإذا لوحظ اختلاف جوهري بينهما، فإنه دليل قاطع على فشل عملية التحديد، وهذا يستدعي إعادة بناء النموذج وتقديره من جديد.أما إذا تشابهت الدالتان، فإننا ننتقل إلى دراسة وتحليل بواقي التقدير مع دالة الارتباط الذاتي للبواقي.

يجب أن تقع معاملات الارتباط الذاتي الكلية لهذه البواقي داخل مجال الثقة المعبر عنه بيانيا بخطين متوازيين $\left[-\frac{t_{\alpha/2}}{\sqrt{T}}, \frac{t_{\alpha/2}}{\sqrt{T}}\right]$.

تحت فرضية التوزيع الطبيعي لدالة الارتباط الذاتي بمتوسط معدوم وتباين $\frac{1}{T}$ أي $\hat{\rho}(k) \sim N\left(0, \frac{1}{T}\right)$، فإن:

$$Q = T\sum_{i=1}^{k} \hat{\rho}^2(i) \sim \chi_\alpha^2(k-p-q)$$

وبمقارنة هذه الإحصائية مع $\chi_\alpha^2(k-p-q)$، نقبل فرضية العدم H_0 إذا كانت Q المحسوبة للأخطاء أقل من تلك المجدولة و هذا يعني أن سلسلة البواقي مستقرة. نشير هنا إلى أنه يمكن استعمال إحصائيةLjung-Box Q^* بدلا من Q:

$$Q^* = T(T+2)\sum_{i=1}^{k} (T-i)\hat{\rho}^2(i) \sim \chi_\alpha^2(k-p-q)$$

عند اختبار الإحصائية Q أو Q^* يمكن رفع مستوى المعنوية من $\alpha = 5\%$ إلى 10%، وهذا الإجراء وارد نظرا لضعف المعنوية في الميدان التطبيقي.

يجب أن تقع أيضا معاملات الارتباط الذاتي الكلية لمربعات البواقي داخل مجال الثقة $\left[-\frac{t_{\alpha/2}}{T},\frac{t_{\alpha/2}}{T}\right]$ ففي هذه الحالة تكون سلسلة مربعات البواقي مستقرة، أي التباين الشرطي للأخطاء متجانس.

2.3.6. اختبار معنوية المعالم والمعنوية الكلية للنموذج:

إذا اعتبرنا أن مقدرات نموذج ARMA(p,q) تتوزع توزيعا طبيعيا، فإن :

$$\frac{\hat{\phi}_i}{\hat{\sigma}_{\hat{\phi}_i}} \sim N(0,1) \quad , \quad i = 1,2,...,p$$

$$\frac{\hat{\theta}_j}{\hat{\sigma}_{\hat{\theta}_j}} \sim N(0,1) \quad , \quad j = 1,2,...,q$$

وهذا المعيار خاص بعملية اختبار المعالم ϕ_i و θ_j:

$$H_0 : \theta_j = 0 \quad , \quad H_0 : \phi_i = 0 \qquad i = 1,2,...,p$$
$$H_1 : \theta_j \neq 0 \quad , \quad H_1 : \phi_i \neq 0 \qquad j = 1,2,...,q$$

نختبر فرضية العدم، حيث نقبل H_0 بمستوى معنوية α إذا كانت $\left|\frac{\hat{\phi}_i}{\hat{\sigma}_{\hat{\phi}_i}}\right| \leq t_{T-p-q,\frac{\alpha}{2}}$، ففي هده الحالة، ليس للمعلم $\phi_i : i = 1,2,...,p$ معنوية إحصائية أي يساوي معنويا الصفر، ونرفض H_0 بمستوى معنوية α إذا كانت $\left|\frac{\hat{\phi}_i}{\hat{\sigma}_{\hat{\phi}_i}}\right| > t_{T-p-q,\frac{\alpha}{2}}$ أي للمعلم ϕ_i معنوية إحصائية أي يختلف معنويا عن الصفر. نفس الشيء بالنسبة لاختبار معنوية أي معلم $\theta_j : j = 1,2,...,q$..

لاختبار المعنوية الكلية للنموذج ARMA(p,q) (غير متضمن لثابتة)، نستخدم إحصائية Fisher. لتكن الفرضيتان:

$$H_0 : \theta_1 = ... = \theta_j = ... = \theta_q = \phi_1 = ... = \phi_i = ... = \phi_p = 0$$

$H_1 : \exists$ معامل $\neq 0$

$$F_c = \frac{\sum_{t=1}^{T}(\hat{Y}_t - \overline{Y})^2 \Big/ (p+q)}{\sum_{t=1}^{T}\hat{\varepsilon}_t^2 \Big/ (T-p-q)} = \frac{R^2 / (p+q)}{(1-R^2)/(T-p-q)} \sim F_\alpha(p+q\,, T-p-q)$$

فإذا تجاوزت الإحصائية F_c قيمة F المجدولة عند مستوى معنوية α ودرجتي حرية $p+q$ و $T-p-q$ نقبل الفرضية القائلة بأن معالم النموذج ليست جميعها مساوية للصفر وأن R^2 يختلف جوهريا عن الصفر. في هذه الحالة، يمكن القول أن للنموذج معنوية إحصائية.

3.3.6. معايير التفضيل بين النماذج المرشحة:

قد يحدث أحيانا في بعض الحالات أن يكون هناك مجموعة من النماذج غير المرفوضة بواسطة الأدوات الإحصائية السابقة الذكر، أي نموذج نختار في هذه الحالة؟ للقيام بعملية المفاضلة بينها نستعمل المعايير التالية:

أ. **معيار Akaike « Akaike Information Criterion » (1969):**

يعد الأكثر استعمالا، ويعطى بالعلاقة التالية:

$$AIC(p,q) = \hat{\sigma}^2.\exp\left\{2\left(\frac{p+q}{T}\right)\right\}$$

حيث $\hat{\sigma}^2$ تباين البواقي المحسوب بطريقة المعقولية العظمى أي بقسمة مربعات البواقي على عدد المشاهدات فقط كما أن المقدار $(p+q)$ هنا يشير إلى عدد معالم النموذج المقدر وليس مجموع درجتي النموذج،كما يمكن كتابة هذا المعيار في شكل لوغاريتمي كما يلي:

$$AIC(p,q) = Ln(\hat{\sigma}^2) + 2\left(\frac{p+q}{T}\right)$$

وبسبب إعطائه وزن أكبر للنماذج المستعملة لأكبر عدد من المشاهدات عُ دِّل كم ا يلي:

$$NAIC(p,q) = \frac{AIC(p,q)}{T}$$

وهنا يكون الاختيار على أساس أصغر قيمة للمعيار، أي نفضل النموذج الذي يحقق أصغر AIC أو NAIC.

ب. معيار (1979) Schwarz « Bayesian Information Criterion »:

رغبةً في تحقيق خصائص تفاربية، اقترح (1979) Schwarz التعديل التالي:

$$BIC = Ln\left(\hat{\sigma}^2\right) + \left(\frac{p+q}{T}\right).LnT$$

يكون أساس اختيار النموذج إذن على أساس أصغر قيمة لهذا المعيار.

ج. معيار Hannan-Quinn (1979):

ويعطى بالعلاقة:

$$HQ(p,q) = Ln\left(\hat{\sigma}^2\right) + (p+q)C^{\frac{Ln\,LnT}{T}} \quad , \quad C > 2$$

حيث $\hat{\sigma}^2$ تباين البواقي المحسوب بطريقة المعقولية العظمى. ويكون النموذج الأفضل حسب هذا المعيار ذلك الذي يعطي أقل قيمة $Min\ HQ(p,q)$.

هناك ملاحظة أخرى تتعلق بإمكانية إضافة متغيرات الانحدار الذاتي والمتوسط المتحرك للنموذج في مرحلة التأكد من التشخيص، ومن ثم ندرس ونختبر معنوياتها الإحصائية، ويمكن أن نستعين في اتخاذ هذا القرار معيار Akaike، كما يمكن اختبار البواقي والنظر ما إذا كانت عشوائية أم لا. إن دوال الارتباط الذاتي للبواقي و مربعاتها يمكن أن تبين ما إذا كان من السهل شرحها بواسطة السيرورة ARMA. فإذا كانت البواقي ممثلة جيدا بواسطة السيرورة AR(p)، يمكن زيادة الدرجة p للسيرورة الأصلية ARMA، وإذا كانت ممثلة جيدا بواسطة السيرورة MA(q)، يمكن زيادة الدرجة q، وبعد إعادة تحديد النموذج،

نعيد تقديره، ونطبق فكرة التأكد من التشخيص مرة أخرى حتى تصبح المعالم $\hat{\theta}_i, \hat{\phi}_j$ ذات معنوية إحصائية والبواقي ذات اضطراب أبيض White Noise.

د. طريقة Goldfrey (1979م) لتشخيص النماذج:

يقترح Godfrey (1979) النموذج التالي:

$$\Phi(L)W_t = \theta(L)\varepsilon_t$$

مع:

$$\Phi(L)= \left(1-\phi_1 L-\phi_2 L^2 -............-\phi_p L^p -\phi_{p+1}L^{p+1} -\phi_{p+2}L^{p+2} -..............-\phi_{p+p^*}L^{p+p^*}\right)$$

$$\theta(L)= \left(1-\theta_1 L-\theta_2 L^2 -............-\theta_q L^q -\theta_{q+1}L^{q+1} -\theta_{q+2}L^{q+2} -..............-\theta_{q+q^*}L^{q+q^*}\right)$$

حيث W_t السلسلة المستقرة بعد إجراء الفروقات من الدرجة d على السلسلة Y_t).

وتكون المعالم θ_{q+j}, ϕ_{p+i} في البداية مساوية للصفر، حيث $i=1,2,....p^*; j=1,2,...q^*$ ثم نقدر النموذج المقترح بالطرق التي تطرقنا إليها سابقا، ومن تم فإن الانحرافات المعيارية للمقدرات المضافة سوف تبين ما إذا كانت هذه المعالم المضافة تختلف عن الصفر أم لا، كما يمكن استعمال اختبار LM المقترح من طرف Godfrey و الذي يعتمد على مشتقات لوغاريتم دالة المعقولية بالنسبة للمعالم المضافة، والمقيمة عند المعالم المقدرة في ظل الفرضية H_0 والقائلة بأن النموذج الأصلي هو الصحيح.

نبين في ظل صحة H_0 بأن لهذه المشتقات توزيعات طبيعية تقاربية، ومن أجل ARMA(p,q) التي تشرح الظاهرة، يمكن كتابة العلاقة:

$$Log\ L\left(\phi,\theta,\sigma_\varepsilon^2\right)=\frac{-T}{2}Log 2\pi-\frac{T}{2}Log\sigma_\varepsilon^2-\sum_{t=1}^{T}\varepsilon_t^2 \Big/ 2\sigma_\varepsilon^2$$

حيث أن:

$$\varepsilon_t = \theta^{-1}(L)\Phi(L)W_t = \left(1+\theta_1 L+\theta_2 L^2+.........+\theta_q L^q\right)^{-1}\left(1+\phi_1 L+\phi_2 L^2+.........+\phi_p L^p\right)W_t$$

وتكون المشتقات الجزئية لهذه الدالة بالنسبة لمعالم الانحدار الذاتي والمتوسط المتحرك هي على الترتيب:

$$\frac{\partial Log L}{\partial \phi_i} = \sum_{t=1}^{T}\left(1+\theta_1 L+\theta_2 L^2+.........+\theta_q L^q\right)^{-1} W_{t-i}.\varepsilon_t \Big/ \sigma_\varepsilon^2$$

$$\frac{\partial Log L}{\partial \theta_j} = \sum_{t=1}^{T}\left(1+\theta_1 L+\theta_2 L^2+.........+\theta_q L^q\right)^{-2}\left(1+\phi_1 L+\phi_2 L^2+.........+\phi_p L^p\right)W_{t-i}.\varepsilon_t \Big/ \sigma_\varepsilon^2$$

$$= \sum_{t=1}^{T}\left(1+\theta_1 L+\theta_2 L^2+.........+\theta_q L^q\right)^{-1}\varepsilon_{t-i}.\varepsilon_t \Big/ \sigma_\varepsilon^2$$

ويعتمد اختبار LM على هذه المشتقات حيث نعتبر النموذج الخاص بالسلسلة ذات الفروقات W_t مع عينة حجمها T.

$$W_t - \phi_1 W_{t-1} - ....- \phi_p W_{t-p} = \varepsilon_t - \theta_1\varepsilon_{t-1} - ....- \theta_1\varepsilon_{t-1}$$

$$\Phi(L)W_t = \theta(L)\varepsilon_t$$

ونمثل مقدرات المعقولية العظمى بواسطة $\hat{\phi}_j$ و $\hat{\theta}_i$ والبواقي بواسطة:

$$\hat{\varepsilon}_t = \hat{\theta}^{-1}(L).\hat{\Phi}(L)\widetilde{W}_t$$

حيث أن $\widetilde{W}_t$ هي القيم الملاحظة فقط للسيرورة W_t، وتكون السلسلتان X_t و Z_t من الشكل:

$$\hat{\theta}(L)X_t = \widetilde{W}_t \Rightarrow X_t = \widetilde{W}_t - \hat{\theta}_1 X_{t-1} - .....- \hat{\theta}_q X_{t-q}$$

$$\hat{\theta}(L)Z_t = \hat{\varepsilon}_t \Rightarrow Z_t = \hat{\varepsilon}_t - \hat{\theta}_1 Z_{t-1} - .....- \hat{\theta}_q Z_{t-q}$$

ونبدأ الحسابات عمليا في المعادلتين السابقتين، بواسطة وضع X_t و Z_t مساوية للصفر من أجل $t = 1-q,.....,-1,0$، ولنعتبر الآن مشكلة اختبار نموذجنا بأنه مخصص بطريقة صحيحة ضد الفرضية البديلة والقائلة بأنه يجب إضافة m معلم في الجزء MA، وبالتالي يجب اختبار الفرضية:

$$H_0: \quad ARMA(p,q)$$
$$H_A: \quad ARMA(p,q+m)$$

ويقترح Godfrey (1979) استعمال اختبار LM عن طريق تقدير نموذج الانحدار التالي بطريقة المربعات الصغرى:

$$\hat{\varepsilon}_t = \alpha_1 X_{t-1} + ...+ \alpha_p X_{t-p} + \beta_1 Z_{t-1} + ...+ \beta_{q+m} Z_{t-q-m} + \mu_t$$

حيث أن α_i و β_j هي معالم، و μ_t هو حد الخطأ، ثم تحت H_0 صحيحة نجري الاختبار:

$$Q = T\left[1 - \frac{\sum \hat{\mu}_t^2}{\sum \hat{\varepsilon}_t^2}\right] \sim \chi_\alpha^2(m)$$

ومن أجل قيم كبيرة لهذه الإحصائية Q نرفض H_0 ، وطوّر Godfrey كذلك إحصائية LM ضد الفرضية البديلة والقائلة بأن النموذج الصحيح هو $ARMA(p+m,q)$ ، حيث نتبع في هذه الحالة نفس الخطوات السابقة.

هـ .. اختبار Granger-Newbold:

يرى Granger and Newbold (1986) أنه بإمكان تطوير اختبار LM لـ . . Goldfrey إلى أي شكل من نوع $ARMA(p+k_1,q+k_2)$ ، حيث أن m هنا تساوي أكبر قيمة بين k_1 و k_2 أي: $m = \max(k_1 k_2)$.

وعلى العموم يجب النظر إلى النموذج الذي يتضمن أصغر عدد من المعالم المتناسقة مع الفرضية القائلة بأن لحدود الأخطاء اضطراب (تشويش) أبيض. كما يمكن في هذا الإطار استعمال المقاييس السابقة الذكر AIC، $NAIC$..

4.6. مرحلة التنبؤ:

إن الهدف من التنبؤ هو استعمال النموذج الحالي والمقدر في فترة زمنية معطاة، من أجل تقدير القيم المستقبلية كسلسلة زمنية تبعا لأصغر خطأ ممكن، لذا نعتبر التنبؤ ذا أصغر متوسط لمربع خطأ التنبؤ Minimum Mean Square Forecast Error (MMSEE) تنبؤا أمثلا، وما دام خطأ التنبؤ متغيرا عشوائيا، نقوم بتصغير قيمته المتوقعة.

إن هذا التنبؤ يتم بعد تقدير معالم النموذج ARIMA(p,d,q)، والذي يكون قد تجاوز مختلف مراحل الاختبارات السابقة ومحددا بالدرجة p، d و q، حيث تصبح قيمة التنبؤ ثابتة (أي تكون مساوية لمتوسط السلسلة) بعد الفترة q في نماذج المتوسطات المتحركة، ويمكن تلخيص عملية التنبؤ في المراحل التالية:

أ- كتابة النموذج المقدر $\hat{Y}_t = f\left(\hat{\phi}, \hat{\theta}, Y_t, \hat{\varepsilon}_t\right)$

ب- تعويض t بـ $T+h$. حيث $h = 1,2,...,H$

ت- تعويض كل القيم المستقبلية للمتغير الخاص بالظاهرة المدروسة بتنبؤاتها، بينما يتم تعويض الأخطاء المستقبلية بالأصفار والماضية (داخل العينة) بالبواقي.

يمكن استعمال النموذج ARIMA المقدر لحساب التنبؤ $\hat{Y}_{T+h}$، حيث نحسب أولا، التنبؤ بفترة واحدة في المستقبل، ثم نستعمل هذا الأخير لحساب التنبؤ بفترتين في المستقبل، ونواصل بنفس الطريقة حتى نصل إلى التنبؤ بالفترة h في المستقبل.ولنكتب نموذج ARIMA(p,d,q) على الشكل:

$$\Phi(L)(1-L)^d Y_t = \delta + \theta(L)\varepsilon_t$$

أو:

$$\Phi(L)\nabla^d Y_t = \delta + \theta(L)\varepsilon_t$$

أو على النحو:

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + ... + \phi_p W_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q} + \delta$$

وهذا يستلزم:

$$\left(1 - \phi_1 L - \phi_2 L^2 - ...... - \phi_p L^p\right) W_t = \delta + \left(1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q\right)\varepsilon_t$$

أي أن السلسلة $W_t = \nabla^d Y_t$ تخضع لنموذج ARMA(p,q):

$$\Phi(L)W_t = \delta + \theta(L)\varepsilon_t$$

ومنه لحساب $\hat{Y}_{T+h}$ نبدأ بحساب تنبؤ W_t من أجل الفترة $T+1$، حيث نستطيع كتابة النموذج في الفترة الزمنية $T+1$:

$$W_{T+1}=\phi_1 W_T+\phi_2 W_{T-1}+......+\phi_p W_{T-p+1}+\varepsilon_{T+1}+\theta_1\varepsilon_T+\theta_2\varepsilon_{T-1}+......+\theta_q\varepsilon_{T-q+1}+\delta$$

ثم نأخذ القيمة المتوقعة الشرطية لـ W_{T+1} . لهدف حساب التنبؤ في الفترة الأولى $\hat{W}_{T+1}$ كما يلي:

$$\hat{W}_{T+1}=E[W_{T+1}\,|\,W_T,.....,W_1]$$
$$=\hat{\phi}_1 W_T+\hat{\phi}_2 W_{T-1}+......+\hat{\phi}_p W_{T-p+1}+\hat{\theta}_1\hat{\varepsilon}_T+\hat{\theta}_2\hat{\varepsilon}_{T-1}+.....+\hat{\theta}_q\varepsilon_{T-q+1}+\hat{\delta}$$

حيث أن $(\hat{\varepsilon}_T,\hat{\varepsilon}_{T-1},..........,\hat{\varepsilon}_{T-q+1})$ هي البواقي المشاهدة، كما أن $E(\varepsilon_{T+1}\,|\,W_T,..)$.
لنستعمل الآن $\hat{W}_{T+1}$ من أجل الحصول على فترة ثانية $\hat{W}_{T+2}$ كما يلي:

$$\hat{W}_{T+2}=E(W_{T+2}\,|\,W_T,W_{T-1}.....,W_1)$$
$$=\hat{\phi}_1\hat{W}_{T+1}+\hat{\phi}_2 W_T+.......+\hat{\phi}_p W_{T-p+2}+\hat{\theta}_1\hat{\varepsilon}_T+.......+\hat{\theta}_q\varepsilon_{T-q+2}+\hat{\delta}$$

ثم نستعمل $\hat{W}_{T+2}$ لنحصل على $\hat{W}_{T+3}$، وهكذا نواصل التعويض إلى أن نصل إلى الحالة العامة:

$$\hat{W}_{T+h}=E[W_{T+h}/W_T,W_{T-1}.....,W_1]$$
$$=\hat{\phi}_1\hat{W}_{T+h-1}+.......+\hat{\phi}_p W_{T+h-p}+\hat{\theta}_1\hat{\varepsilon}_{T+h-1}+.......+\hat{\theta}_q\varepsilon_{T+h-q}+\hat{\delta}$$

لدراسة دقة التنبؤ الذي يعتبر من أهم المراحل في تقييم النموذج للأغراض المستقبلية، نستخدم في هذا المجال متوسط الخطأ الذي يعبر على متوسط الفرق بين المشاهدة والتنبؤ لنفس الفترة الزمنية، ويُعطى رياضيا في الشكل التالي:

$$MRAE=H^{-1}\sum_{h=1}^{H}\frac{\left|\hat{Y}_{T+h}-Y_{T+h}\right|}{\left|Y_{T+h}\right|}\times 100$$

حيث Y_{T+h} و $\hat{Y}_{T+h}$ تعبران عن السلسلة المدروسة المتنبأ بها نظريا و تلك المتنبأ بها تقديريا على الترتيب. ويمكن أن يؤخذ هذا المقياس في شكل نسبي وكما يلي:

$$PME=H^{-1}\sum_{h=1}^{H}\left(\frac{Y_{T+h}-\hat{Y}_{t+h}}{Y_{T+h}}\right)$$

يمكن أيضا استخدام متوسط مربع الخطأ الذي يعتبر أكثر فعالية من العيار السابق، لدينا:

$$QME = H^{-1}\sum_{h=1}^{H}(\hat{Y}_{T+h} - Y_{T+h})^2$$

حيث H هي عدد القيم المتوقعة مع $h = 1,2,...H$.

يستخدم بعض الإحصائيين معيارا آخر يسمى بمعيار ثايل Theil's U statistic: وهو معطى بالعلاقة التالية:

$$U = \frac{\sqrt{H^{-1}\sum_{h=1}^{H}(\hat{Y}_{T+h} - Y_{T+h})^2}}{\sqrt{H^{-1}\sum_{h=1}^{H}Y_{T+h}^2} + \sqrt{H^{-1}\sum_{h=1}^{H}\hat{Y}_{T+h}^2}}$$

ويكون التنبؤ جيدا عندما يكون $U = 0$، وتكون العملية فاشلة عندما $U = 1$، وعمليا يتذبذب هذا المقياس بين هاتين القيمتين.

يمكن أيضا قياس دقة التنبؤ من خلال مدى قدرة التنبؤ في اقتفاء أثر السلسلة الأصـلية والقدرة على تتبع نقاط انعطافها برشاقة كما ذكرنا سابقا، ولتوضيح هذه العملية نستعين دائما بالرسومات البيانية للسلسلتين الأصلية والتنبؤية.

مثال 5:

وفقا للنتائج المتحصل عليها في المثال 4، نقوم بتقدير النموذجين بطريقة Gauss-Newton، ففي هذه الحالة يمكن استخدام برمجيتي RATS 5.04 أو Eviews 5.0. للمفاضلة بين النموذجين يمكن استعمال برمجية Eviews من أجل حساب معياري AIC و Schwarz لكل نموذج ثم نستعمل RATS 5.04 لأنه يعطي نتائج أحسن لقيمة معامل التحديد. تظهر النتائج على النحو التالي:

- بالنسبة للنموذج AR(1):

Dependent Variable: DLOGLQT
Method: Least Squares
Date: 01/01/03 Time: 01:10
Sample(adjusted): 1972 2008
Included observations: 37 after adjusting endpoints
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.071838	0.015091	4.760274	0.0000
AR(1)	-0.472218	0.085987	-5.491744	0.0000

R-squared	0.462854	Mean dependent var	0.062003
Adjusted R-squared	0.447507	S.D. dependent var	0.181224
S.E. of regression	0.134704	Akaike info criterion	-1.118942
Sum squared resid	0.635077	Schwarz criterion	-1.031866
Log likelihood	22.70044	F-statistic	30.15925
Durbin-Watson stat	0.535745	Prob(F-statistic)	0.000004
Inverted AR Roots	-.47		

- بالنسبة للنموذج MA(1):

Dependent Variable: DLOGLQT
Method: Least Squares
Date: 01/01/03 Time: 01:41
Sample(adjusted): 1971 2008
Included observations: 38 after adjusting endpoints
Convergence achieved after 7 iterations
Backcast: 1970

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.084791	0.010360	8.184140	0.0000
MA(1)	-0.704057	0.120575	-5.839154	0.0000

R-squared	0.424908	Mean dependent var	0.092084
Adjusted R-squared	0.408933	S.D. dependent var	0.257565
S.E. of regression	0.198018	Akaike info criterion	-0.349717
Sum squared resid	1.411607	Schwarz criterion	-0.263529
Log likelihood	8.644632	F-statistic	26.59868
Durbin-Watson stat	1.339223	Prob(F-statistic)	0.000009
Inverted MA Roots	.70		

نلاحظ أن النموذج الأمثل الذي يعبر أكثر عن تغيرات سلسلة الطلب على الكهرباء هو نموذج ARIMA(0,1,1)بالرغم من أن معياري AIC و Schwarz يشيران إلى أفضلية ARIMA(1,1,0). السبب الذي جعلنا نختار نموذج المتوسط المتحرك هو جودة إحصائية دربين–واتسون عكس النموذج AR(1) الذي يظهر ارتباطا ذاتيا بين الأخطاء.

من الملاحظ أن معامل التحديد المتحصل عليه ليس مرتفعا جدا في كلا النموذجين، فإذا قمنا بتقدير النموذج المختار باستعمال برمجية RATS 5.04، نتحصل على معامل تحديد مرتفع جدا، لدينا:

```
boxjenk(constant,ar=0,ma=1,diff=1) logy / resids

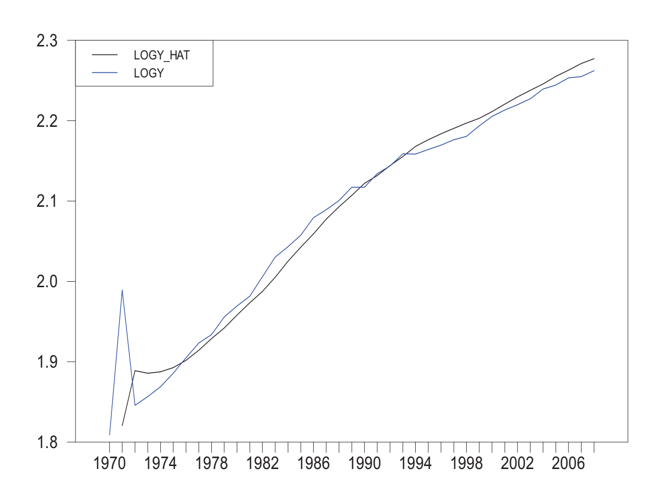
Box-Jenkins - Estimation by Gauss-Newton
Convergence in    13 Iterations. Final criterion was  0.0000037 <  0.0000100
Dependent Variable LOGY
Annual Data From 1971:01 To 2008:01
Usable Observations     38     Degrees of Freedom    36
Centered R**2     0.940143      R Bar **2   0.938480
Uncentered R**2   0.999782      T x R**2      37.992
Mean of Dependent Variable      2.0883104723
Std Error of Dependent Variable 0.1280847331
Standard Error of Estimate      0.0317691918
Sum of Squared Residuals        0.0363341357
Durbin-Watson Statistic             1.281192
Q(9-1)                              0.640929
Significance Level of Q           0.99965946

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  CONSTANT                  0.011454527  0.001847147      6.20120  0.00000037
2.  MA{1}                    -0.661970589  0.129426132     -5.11466  0.00001058
```

نقوم الآن بتشخيص النموذج. نلاحظ أن للمعالم معنوية إحصائية بنسبة معنوية 0.05 باعتبار أن قيم ستيودنت بالقيمة المطلقة أكبر تماما من القيمة الحرجة للتوزيع الطبيعي، إضافة إلى ذلك، للنموذج قدرة تفسيرية عالية جدا.

من خلال الشكل أدناه يمكننا ملاحظة شبه المطابقة بين منحنيي السلسلة الأصلية Actual ومنحنى السلسلة المقدرة Fitted، هذا من شأنه أن يعطينا فكرة عن مدى أهمية تعبير النموذج المقدر ARIMA(0,1,1) إلى بيانات الطلب على الكهرباء.

الشكل (11): السلسلة الأصلية و السلسلة المقدرة



من خلال الشكل (18)، نستنتج أن سلسلة البواقي مستقرة حيث أن معاملات الارتباط الذاتي تقع كلها داخل مجال الثقة $\left[\frac{-1.96}{\sqrt{T}},\frac{+1.96}{\sqrt{T}}\right]$ و هذا يعني أن هناك استقلالية تامة بين الأخطاء. يمكن التأكد من ذلك باستعمال إحصائية Ljung-Box التي تساوي 1.914 تبقى دائما أقل من القيمة المجدولة لتوزيع χ^2 بدرجة حرية 16. كما أن معاملات الارتباط الذاتي لسلسلة مربعات البواقي المبينة في الشكل (19) تساوي معنويا الصفر (تقع كلها داخل مجال الثقة) و هذا يعني أن الأخطاء العشوائية تتميز بتباين شرطي ثابت (متجانس).

الشكل (12): معاملات الارتباط الذاتي و الجزئي للبواقي

Sample: 1971 2008
Included observations: 38
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.027	0.027	0.0304	
		2	0.017	0.016	0.0420	0.838
		3	0.000	0.000	0.0420	0.979
		4	0.003	0.002	0.0424	0.998
		5	0.011	0.010	0.0476	1.000
		6	0.017	0.017	0.0616	1.000
		7	-0.036	-0.037	0.1237	1.000
		8	-0.020	-0.019	0.1445	1.000
		9	-0.053	-0.051	0.2927	1.000
		10	-0.105	-0.103	0.8940	1.000
		11	-0.078	-0.073	1.2352	1.000
		12	-0.034	-0.029	1.3026	1.000
		13	-0.059	-0.056	1.5134	1.000
		14	-0.073	-0.072	1.8479	1.000
		15	-0.015	-0.011	1.8626	1.000
		16	-0.027	-0.027	1.9140	1.000

الشكل (13): معاملات الارتباط الذاتي و الجزئي لمربعات البواقي

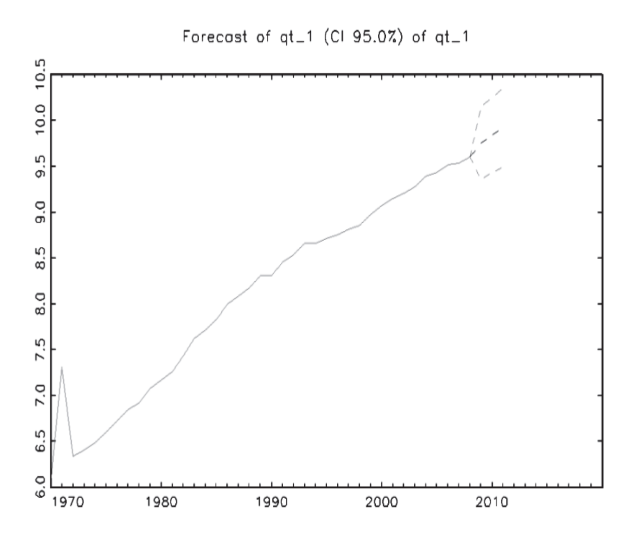
Sample: 1971 2008
Included observations: 38
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.200	0.200	1.6377	
		2	0.102	0.065	2.0781	0.149
		3	0.045	0.013	2.1651	0.339
		4	0.004	-0.014	2.1656	0.539
		5	-0.015	-0.019	2.1767	0.703
		6	-0.020	-0.013	2.1949	0.822
		7	-0.021	-0.012	2.2158	0.899
		8	-0.017	-0.008	2.2310	0.946
		9	-0.017	-0.010	2.2465	0.973
		10	-0.016	-0.009	2.2602	0.987
		11	0.001	0.008	2.2603	0.994
		12	0.020	0.021	2.2841	0.997
		13	0.004	-0.005	2.2851	0.999
		14	-0.001	-0.006	2.2851	1.000
		15	0.011	0.010	2.2931	1.000
		16	-0.011	-0.015	2.3013	1.000

بما أن النموذج مقبول إحصائيا، يمكن إذن التنبؤ بالطلب على الكهرباء في الجزائر على المدى القصير، لنأخذ مثلا ثلاث سنوات (من 2009 إلى غاية 2011). نستعين ببرمجية GAUSS 6.0، النتائج تظهر على الشكل التالي:

time	lower CI	forecast	upper CI	std. err
2009	9.3522	9.7503	10.1484	0.2031
2010	9.4222	9.8369	10.2516	0.2116
2011	9.4929	9.9235	10.3540	0.2197

الشكل (14): التنبؤ و مجالات الثقة للتنبؤ



بعد حساب التنبؤ النقطي يجب دوما بناء فترات ثقة لهذا الأخير لكي يكون التحليل دقيقا بغية اتخاذ القرارات الاقتصادية. من خلال الشكل أعلاه، يمكن القول أن التنبؤ يتبع السلسلة الأصلية مما يؤكد مرة أخرى على الجودة الإحصائية للنموذج المختار و أيضا على قوة التنبؤ.

الفصل السابع

مدخل إلى نماذج VAR ومشكل التكامل المشترك

الفصل السابع

مدخل إلى نماذج VAR ومشكل التكامل المشترك

تطرقنا في الفصل الخامس إلى نماذج المعادلات الآنية التي عرفت انتقادات كثيرة ولاسيما ضعف التنبؤات الناتجة عنها في ظل بيئة اقتصادية معكرة. جاءت نماذج VAR كبديل لهذا النوع من النماذج التنبؤية، فلقد أثبتت الاختلالات الاقتصادية (الأزمة الاقتصادية العالمية،...الخ) عدم صلاحيتها بسبب آنية العلاقات التي تربط بين المتغيرات الاقتصادية وعدم أخذه بعين الاعتبار ديناميكية (حركية) نظام المعادلات القياسية. في نماذج VAR، تُعالج كل المتغيرات بصفة متماثلة و بدون شرط إقصاء مع إدخال عامل التباطؤ لكل المتغيرات في كل المعادلات ليعطي للنظام الطبيعة الحركية. هذه النماذج عبارة عن تعميم لنماذج الانحدار الذاتي إذ يتكون من نظام لجملة معادلات بحيث كل متغيرة هي عبارة عن توليفة خطية لقيمها الماضية و القيم الماضية لمتغيرات أخرى بالإضافة إلى الأخطاء العشوائية. سنتطرق في هذا الفصل إلى نماذج شعاع الانحدار الذاتي و طرق تقديرها ونقوم بإعطاء بعض اختبارات السببية وطرق تحليل الصدمات. في الجزء الثاني من هذا الفصل ستكون الدراسة منصبة حول مشكل التكامل المشترك الذي يسمح بتحديد العلاقة الحقيقية الموجودة بين متغيرين والذي يعتبر بمثابة مفهوم جديد في ميدان الاقتصاد القياسي المطبق على السلاسل الزمنية حيث نعطي بعض الاختبارات التي تساعدنا على كشف التكامل المشترك.

1. نماذج الانحدار الذاتي المتعدد Multivariate Autoregressive models

1.1. الصياغة العامة لنموذج VAR (VARMA):

يكتب نموذج VAR "Vector AutoRegressive" لـ k متغير و p تباطؤ على الشكل المصفوفي التالي:

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p} + \varepsilon_t \quad t = 1,2,...,T \text{ ,}$$

مع:
$$Y_t = \begin{pmatrix} Y_{1,t} \\ Y_{2,t} \\ . \\ . \\ . \\ Y_{k,t} \end{pmatrix} \; ; \; \Phi_i = \begin{pmatrix} \phi_{1i}^1 & \phi_{1i}^2 & .... & \phi_{1i}^k \\ \phi_{2i}^1 & \phi_{2i}^2 & .... & \phi_{2i}^k \\ . & . & & . \\ . & . & .... & . \\ . & . & & . \\ \phi_{ki}^1 & \phi_{ki}^2 & .... & \phi_{ki}^k \end{pmatrix} \; ; \; \Phi_0 = \begin{pmatrix} \phi_1^0 \\ \phi_2^0 \\ . \\ . \\ . \\ \phi_k^0 \end{pmatrix} \; ;$$

$$\varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ . \\ . \\ . \\ \varepsilon_{kt} \end{pmatrix} \text{ ، } i = 1,2,..., p$$

نسمي $\Sigma_\varepsilon = E(\varepsilon_t \varepsilon_t^{'})$ مصفوفة التباين–التباين المشترك للأخطاء وهي ذات بعد (k,k). يمكن أيضا كتابة النموذج بدلالة معامل التأخير حيث:

$$\left(I - \Phi_1 L - \Phi_2 L^2 - ..... - \Phi_p L^p\right) Y_t = \Phi_0 + \varepsilon_t$$

أو:

$$\Phi(L) Y_t = \Phi_0 + \varepsilon_t$$

المتغيرات $Y_{1,t},....,Y_{k,t}$ تعتبر كسلاسل مستقرة والأخطاء $\varepsilon_{1t},......,\varepsilon_{kt}$ ذات تشويش أبيض مستقلة ذاتيا وذات تباينات ثابتة $\sigma_{\varepsilon 1}^2,......,\sigma_{\varepsilon k}^2$.

تكون السيرورة VAR مستقرة إذا وفقط إذا تحققت الفرضيات الكلاسيكية الثلاثة:

- $E(Y_t) = \mu \; , \; \forall t$
- $\mathrm{var}(Y_t) < \infty$
- $\mathrm{cov}(Y_t, Y_{t+k}) = E\left[(Y_t - \mu)(Y_{t+k} - \mu)\right] = \Gamma(k) \; , \; \forall t$

بصفة عامة، تكون السيرورة VAR مستقرة إذا كان كثير الحدود المعرف انطلاقا من محدد المصفوفة $\left|I - \Phi_1 L - \Phi_2 L^2 - ..... - \Phi_p L^p\right| = 0$ تحتوي على جذور خارج الدائرة الوحدوية.

مثال 1:

لتكن السيرورة:

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \begin{pmatrix} 0.2 & 0.7 \\ 0.3 & 0.4 \end{pmatrix} \begin{pmatrix} Y_{1t-1} \\ Y_{2t-1} \end{pmatrix} + \begin{pmatrix} \hat{\varepsilon}_{1t} \\ \hat{\varepsilon}_{2t} \end{pmatrix}$$

نقوم بدراسة شروط استقرارية هذا النموذج. نحسب المحدد التالي:

$$\left| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.2 & 0.7 \\ 0.3 & 0.4 \end{pmatrix} L \right| = 0 \Rightarrow 1 - 0.6L - 0.13L^2 = 0 \Rightarrow \begin{cases} L_1 = 1.30 \\ L_2 = -5.91 \end{cases}$$

نلاحظ أن الجذرين المتحصل عليهما بالقيمة المطلقة أكبر تماما من الواحد وهذا يعني أن النموذج مستقر.

يمكن تعميم نموذج VAR إلى نموذج يحتوي على أخطاء مرتبطة ذاتيا من الدرجة (الرتبة) q:

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p} + \Theta_1 \varepsilon_{t-1} + \Theta_2 \varepsilon_{t-2} + ... + \Theta_q \varepsilon_{t-q} + \varepsilon_t$$

وهو نموذج ARMA(p,q) متعدد المتغيرات أو VARMA(p,q) الذي يصطلح على تسميته أيضا بـ . ARMAX(p,q). تكون السيرورة VMA دائما مستقرة وقابلة للقلب إذا كانت جذور كثير الحدود تقع كلها خارج الدائرة الوحدوية.

قد يتضمن نموذج VAR متغيرات خارجية (مستقلة) ويسمى بنموذج SVAR "Structural Vector AutoRegressive" الذي يأخذ الشكل التالي:

$$, \quad t = 1,2,...,T$$

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p} + B_1 X_{t-1} + ... + B_m X_{t-m} + \varepsilon_t$$

حيث $Y_{1,t},....,Y_{k,t}$ تعبر عن المتغيرات الداخلية و $X_{1t}, X_{2t},....,X_{rt}$ متغيرات خارجية يمكن أن تحتوي على مركبات عشوائية أو غير عشوائية. على العموم يطلق على هذا النموذج باسم النظام الخطي لوجود العلاقة الخطية بين كل المتغيرات ويسمى أيضا بنموذج المعادلات الآنية الحركية (الديناميكية). باستعمال معامل التباطؤ في النموذج، يكون الشكل المختصر كما يلي:

$$\Phi(L)Y_t = B(L)X_t + \varepsilon_t$$

وبضرب الشكل المختصر بـ . $\Phi^{-1}(L)$:

$$Y_t = \Phi^{-1}(L)B(L)X_t + \Phi^{-1}(L)\varepsilon_t$$

حيث $A(L) = \Phi^{-1}(L)B(L)$ يسمى بالشكل النهائي للنظام ويكون هذا الشكل موجودا في حالة ما إذا كانت المصفوفة $\Phi(L)$ قابلة للقلب تحت الشرط التالي:

$$\det(\Phi(L)) \neq 0$$

2.1. تحديد وتقدير نموذج VAR:

في حالة النموذج VAR، يمكن تقدير كل معادلة من معادلات هذا النموذج بطريقة المربعات الصغرى أو بطريقة المعقولية العظمى. يتم تقدير كل معادلة على حدا. النموذج VAR(p) المقدر يكتب على الشكل التالي:

$$\hat{Y}_t = \hat{\Phi}_0 + \hat{\Phi}_1 Y_{t-1} + \hat{\Phi}_2 Y_{t-2} + ... + \hat{\Phi}_p Y_{t-p}$$

نسمي $\Sigma_{\hat{\varepsilon}}$ مصفوفة التباين-التباين المشترك لبواقي التقدير.

لا يمكن تقدير معاملات هذا النموذج انطلاقا من سلاسل غير مستقرة. إذن يجب جعل كل السلاسل مستقرة بحساب الفروقات من الدرجة d في حالة اتجاه عام عشوائي أو إضافة مركبة الاتجاه العام إلى صيغة النموذج VAR في حالة اتجاه عام ثابت. أيضا، يمكن إضافة متغيرات صورية لتصحيح التغيرات الموسمية.

لتحديد درجة النموذج VAR، نستخدم معايير المعلومات، فطريقة اختيار الدرجة تكمن في تقدير كل معادلات النموذج من أجل أي رتبة (درجة) من 0 إلى P (P هو العدد الأقصى المقبول من طرف النظرية الاقتصادية). نستعمل مثلا المعايير الثلاثة Akaike و Hannan-Quin و Schwarz المعرفة كما يلي:

$$AIC = \ln|\Sigma_{\hat{\varepsilon}}| + \frac{2k^2 p}{T}$$

$$HQ = \ln|\Sigma_{\hat{\varepsilon}}| + \frac{2\log\log T}{T} k^2 p$$

$$SC = \ln|\Sigma_{\hat{\varepsilon}}| + \frac{k^2 p \ln(T)}{T}$$

مع k عدد متغيرات النظام، T عدد المشاهدات، p عدد الفجوات الزمنية، $\Sigma_{\hat{\varepsilon}}$ مصفوفة التباين-التباين المشترك للبواقي.

نختار التباطؤ الأمثل وذلك بتصغير المعايير الثلاثة. يمكن أيضا استخدام نسبة المعقولية لهذا الغرض انطلاقا من تقدير تباين البواقي. إذا كان $\Sigma^1_{\hat{\varepsilon}}$ تباين بواقي النموذج المقيد و $\Sigma^0_{\hat{\varepsilon}}$ تباين النموذج الأول (غير المقيد)، فإن إحصائية نسبة المعقولية $T\left(\ln\left|\Sigma^1_{\hat{\varepsilon}}\right|-\ln\left|\Sigma^0_{\hat{\varepsilon}}\right|\right)$ تتوزع توزيع χ^2 بدرجة حرية تساوي عدد القيود.

3.1. التنبؤ:

بعد تقدير معاملات النظام ، يتم حساب التنبؤ في الفترة T من أجل $T+1$ لنموذج VAR(1):

$$\hat{Y}_T(1)=\hat{\Phi}_0+\hat{\Phi}_1 Y_T$$

من أجل $T+2$، يكون التنبؤ محسوبا كما يلي:

$$\hat{Y}_T(2)=\hat{\Phi}_0+\hat{\Phi}_1\hat{Y}_T(1)=\hat{\Phi}_0+\hat{\Phi}_1\hat{\Phi}_0+\hat{\Phi}_1^2 Y_T$$

من أجل $T+3$، لدينا:

$$\hat{Y}_T(3)=\hat{\Phi}_0+\hat{\Phi}_1\hat{Y}_T(2)=\left(I+\hat{\Phi}_1+\hat{\Phi}_1^2\right)\hat{\Phi}_0+\hat{\Phi}_1^3 Y_T$$

من أجل $T+h$ (h هو أفق التنبؤ)، يكون التنبؤ على الشكل التالي:

$$\hat{Y}_T(h)=\left(I+\hat{\Phi}_1+\hat{\Phi}_1^2+....+\hat{\Phi}_1^{h-1}\right)\hat{\Phi}_0+\hat{\Phi}_1^h Y_T$$

عندما يؤول h إلى ما لا نهاية، فالتنبؤ يؤول إلى حالة مستقرة لأن $\hat{\Phi}_1^h\to 0$ إذا كان $h\to\infty$ وخطأ التنبؤ $\hat{\varepsilon}_{T+h}=Y_{T+h}-\hat{Y}_T(h)$ متوسطه معدوم و تباينه معطى بالعلاقة:

$$\Sigma_{\hat{\varepsilon}_{T+h}}=M_0\Sigma_{\hat{\varepsilon}}M_0^{'}+M_1\Sigma_{\hat{\varepsilon}}M_1^{'}+.....+M_{h-1}\Sigma_{\hat{\varepsilon}}M_{h-1}^{'}$$

حيث M_i محسوبة بصيغة التراجع:

$$,\quad i=1,2,......\ M_i=\sum_{j=1}^{\min(p,i)}\Phi_j M_{i-j}$$

و:

$$M_0=I$$

لدينا:

$$M_1=\hat{\Phi}_1;M_2=\hat{\Phi}_1 M_1+\hat{\Phi}_2 M_0=\hat{\Phi}_1^2+\hat{\Phi}_2\ ;$$

$$M_2 = \hat{\Phi}_1 M_1 + \hat{\Phi}_2 M_0 + \hat{\Phi}_3 M_0 = \hat{\Phi}_1^2 + \hat{\Phi}_1\hat{\Phi}_2 + \hat{\Phi}_2\hat{\Phi}_1 + \hat{\Phi}_3$$

تباين خطأ التنبؤ لكل متغير يتم قراءته في قطر المصفوفة $\Sigma_{\hat{\varepsilon}_{T+h}}$ ومجال التنبؤ عند مستوى $1-\alpha/2$ معطى بـ :.

$$Y_{T+h} \in \left[\hat{Y}_T(h) - z_{\alpha/2}\ \text{var}(\hat{\varepsilon}_{T+h}), \hat{Y}_T(h) + z_{\alpha/2}\ \text{var}(\hat{\varepsilon}_{T+h})\right]$$

حيث $z_{\alpha/2}$ هي القيمة الحرجة للتوزيع الطبيعي.

مثال 2:

نريد نمذجة الناتج الوطني الإجمالي، Y_{1t} و المخزون النقدي Y_{2t} باستعمال نموذج VAR. المعطيات فصلية من 1965 إلى 2005. ليكن النموذج:

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} \phi_1^0 \\ \phi_2^0 \end{pmatrix} + \begin{pmatrix} \phi_{11}^1 & \phi_{11}^2 \\ \phi_{21}^1 & \phi_{22}^2 \end{pmatrix} \begin{pmatrix} Y_{1t-1} \\ Y_{2t-2} \end{pmatrix} + \ldots + \begin{pmatrix} \phi_{14}^1 & \phi_{14}^2 \\ \phi_{24}^1 & \phi_{24}^2 \end{pmatrix} \begin{pmatrix} Y_{1t-4} \\ Y_{2t-4} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

لتقدير معالم النموذج VAR، نطبق طريقة المربعات الصغرى العادية على كل معادلة. نختار النموذج الأمثل بالاستعانة بالمعيارين Akaike و Schwarz من أجل رتبة (درجة) تتغير من 0 إلى 4. نقدر أربع نماذج مختلفة ونختار النموذج الذي يصغر المعيارين AIC و Schwarz، نستعمل لهذا الغرض برمجية RATS 5.04:

```
DISPLAY
DO LAGS=1,4
  SYSTEM
  VARIABLES Y1 Y2
  LAGS 1 TO LAGS
  DET CONSTANT
  END(SYSTEM)
  ESTIMATE (NOPRINT, SIGMA)
  COMPUTE SCHWARZ=%LOGDET+(4*LAGS*LOG(%NOBS))/%NOBS
  COMPUTE AIC=%LOGDET+(2*4*LAGS)/%NOBS
  IF LAGS==1
    DISPLAY @4'LAGS'  @20 'AIC'  @35  'SCHWARZ'
  DISPLAY @5 #### LAGS  @20 ####.#### AIC  @35 SCHWARZ
END DO LAGS
```

LAGS	AIC	SCHWARZ
1	-15.41	-15.86
2	-15.22	- 15.56
3	-14.87	-15.03
4	-14.75	-14.84

نلاحظ أن المعيارين AIC و Schwarz يأخذان قيما صغرى عند $p=1$. البرنامج الذي يسمح بتقدير النموذج الأمثل هو كالآتي:

```
COMPUTE LAGS=1
SYSTEM
VARIABLES Y1 Y2
LAGS 1 TO LAGS
DET CONSTANT
END(SYSTEM)
ESTIMATE (SIGMA, RESIDS=RESIDS2)
```

النموذج VAR(1) المقدر يكتب على الشكل التالي:

$$\hat{Y}_{1t} = 0.096 + 0.28Y_{1t-1} + 0.005Y_{2t-1}, R^2 = 0.99$$
$$(1.76) \quad (50.93) \quad (1.06)$$

$$\hat{Y}_{2t} = -0.166 + 0.032Y_{1t-1} + 0.29Y_{2t-1}, R^2 = 0.99$$
$$(-2.55) \quad (2.75) \quad (70.54)$$

(.):قيم ستيودنت.

بواقي التقدير $\hat{\varepsilon}_{1t}$ و $\hat{\varepsilon}_{2t}$ محسوبة مع مصفوفة التباين-التباين المشترك للبواقي:

$\Sigma_{\hat{\varepsilon}_t} = \begin{pmatrix} 109.54 & 45.13 \\ 45.13 & 101.37 \end{pmatrix}$. مع:

$$\text{var}(\hat{\varepsilon}_{1t}) = 109.54 \ ; \ \text{var}(\hat{\varepsilon}_{2t}) = 101.37 \ ; \ \text{cov}(\hat{\varepsilon}_{1t}, \varepsilon_{2t}) = 45.13$$

لحساب التنبؤ، نستعين بالنموذج المقدر:

$$\hat{Y}_{1T}(1) = 0.096 + 0.28Y_{1T} + 0.005Y_{2T}$$
$$\hat{Y}_{2T}(1) = \text{-}0.166 + 0.032Y_{1T} + 0.29Y_{2T}$$

$$\hat{Y}_{1T}(2) = 0.096 + 0.28\hat{Y}_{1T}(1) + 0.005\hat{Y}_{2T}(1)$$
$$\hat{Y}_{2T}(2) = \text{-}0.166 + 0.032\hat{Y}_{1T}(1) + 0.29\hat{Y}_{2T}(1)$$

$$\hat{Y}_{1T}(3) = 0.096 + 0.28\hat{Y}_{1T}(2) + 0.005\hat{Y}_{2T}(2)$$
$$\hat{Y}_{2T}(3) = \text{-}0.166 + 0.032\hat{Y}_{1T}(2) + 0.29\hat{Y}_{2T}(2)$$

$$\hat{Y}_{1T}(4) = 0.096 + 0.28\hat{Y}_{1T}(3) + 0.005\hat{Y}_{2T}(3)$$
$$\hat{Y}_{2T}(4) = \text{-}0.166 + 0.032\hat{Y}_{1T}(3) + 0.29\hat{Y}_{2T}(3)$$

من أجل $h = 1$، لدينا $\Sigma_{\hat{\varepsilon}_{T+1}} = \begin{pmatrix} 109.56 & 45.14 \\ 45.14 & 101.38 \end{pmatrix}$. تباين خطأ التباين لـ $\hat{Y}_{1T}(1)$ يساوي 109.56 وتباين خطأ التنبؤ لـ $\hat{Y}_{2T}(1)$ يساوي 101.38. بعد استنتاج قيم الخطأين المعياريين للتنبؤ، يمكن بناء فترات ثقة للتنبؤ بنسبة احتمال 95%.

2. التحليل الهيكلي Structural Analysis

1.2. السببية Causality:

يعتبر مشكل السببية من أهم المحاور في تحديد صيغ النماذج الاقتصادية، إذ يهدف إلى البحث عن أسباب الظواهر الاقتصادية وفهمها للتمييز بين الظاهرة التابعة من الظواهر المستقلة المُفَسِرة لها.

1.1.2. اختبار السببية وفق Granger:

اقترح Granger (1969) معيار تحديد العلاقة السببية التي ترتكز على العلاقة الديناميكية الموجودة بين السلاسل الزمنية، حيث إذا كانت Y_{1t} و Y_{2t} سلسلتين زمنيتين

تعبران عن تطور ظاهرتين اقتصاديتين مختلفتين عبر الزمن t، وكانت السلسلة Y_{1t} تحتوي على المعلومات التي من خلالها يمكن تحسين التوقعات بالنسبة للسلسلة Y_{2t}، في هذه الحالة نقول أن Y_{1t} تُسَبِّب Y_{2t}، إذن نقول عن متغيرة أنها سببية إذا كانت تحتوي على معلومات تساعد على تحسين التوقع لمتغيرة أخرى.

يُستخدم اختبار Granger في التأكد من مدى وجود علاقة تغذية مرتدة أو استرجاعية Feedback أو علاقة تبادلية بين متغيرين، وذلك في حالة وجود بيانات سلسلة زمنية.

ومن المشاكل التي توجد في هذه الحالة أن بيانات السلسلة الزمنية لمتغير ما كثيرا ما تكون مرتبطة، أي يوجد ارتباط ذاتي بين قيم المتغير الواحد عبر الزمن، ولاستبعاد أثر هذا الارتباط الذاتي إن وجد، يتم إدراج قيم نفس المتغير التابع لعدد من الفجوات الزمنية كمتغيرات تفسيرية في علاقة السببية المراد قياسها، يُضاف إلى ذلك إدراج قيم المتغير التفسيري الآخر لعدد من الفجوات الزمنية كمتغيرات تفسيرية أيضا، وذلك باعتبار أن السبب يسبق النتيجة في الزمن.

ليكن النموذج VAR(p) المستقر حيث:

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} \phi_1^0 \\ \phi_2^0 \end{pmatrix} + \begin{pmatrix} \phi_{11}^1 & \phi_{11}^2 \\ \phi_{21}^1 & \phi_{21}^2 \end{pmatrix} \begin{pmatrix} Y_{1t-1} \\ Y_{2t-1} \end{pmatrix} + \begin{pmatrix} \phi_{12}^1 & \phi_{12}^2 \\ \phi_{22}^1 & \phi_{22}^2 \end{pmatrix} \begin{pmatrix} Y_{1t-2} \\ Y_{2t-2} \end{pmatrix} + \ldots + \begin{pmatrix} \phi_{1p}^1 & \phi_{1p}^2 \\ \phi_{2p}^1 & \phi_{2p}^2 \end{pmatrix} \begin{pmatrix} Y_{1t-p} \\ Y_{2t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

السلاسل $Y_{2t-1}, Y_{2t-2}, \ldots\ldots, Y_{2t-p}$ تعتبر كمتغيرات خارجية بالنسبة للمتغيرات $Y_{1t-1}, Y_{1t-2}, \ldots\ldots, Y_{1t-p}$، نرى ما إذا كانت مجموعة Y_{2t} لا تحسن معنويا من القدرة التفسيرية للمتغيرات Y_{1t} للنموذج VAR والذي نطلق عليه تسمية RVAR " Restricted VAR". اختيار الفجوات الزمنية يتم بواسطة المعيارين AIC و Schwarz، ليكن:

- Y_{2t} لا يسبب Y_{1t} إذا كانت الفرضية $H_0 : \phi_{11}^2 = \phi_{12}^2 = \ldots = \phi_{1p}^2 = 0$ مقبولة
- Y_{1t} لا يسبب Y_{2t} إذا كانت الفرضية $H_0 : \phi_{21}^1 = \phi_{22}^1 = \ldots = \phi_{2p}^1 = 0$ مقبولة

إذا قبلنا الفرضيتين معا، نتحدث هنا عن ما يسمى بـ "Feed Back effect". يمكن استعمال إحصائية فيشر للقيام بالاختبار و هو اختبار انعدام المعاملات، معادلة بمعادلة أو مباشرة المقارنة بين نموذج VAR غير المقيد UVAR والنموذج VAR المقيد RVAR.

نحسب نسبة المعقولية $L^* = (T-c)\times\left(\ln|\Sigma_{RVAR}| - \ln|\Sigma_{UVAR}|\right)$ التي تتبع توزيع χ^2 بدرجة حرية $2\times p$ مع:

Σ_{RVAR}: مصفوفة التباين-التباين المشترك لبواقي النموذج المقيد،

Σ_{UVAR}: مصفوفة التباين-التباين المشترك لبواقي النموذج غير المقيد،

T: عدد المشاهدات،

c: عدد المعالم المقدرة في كل معادلة للنموذج غير المقيد.

إذا كانت $L^* > \chi^2_{\alpha}(2p)$، ففي هذه الحالة نرفض فرضية وجود القيود، أي هناك سببية وفق Granger.

2.1.2. اختبار السببية وفق Sims:

اقترح Sims (1980) اختبارا آخر مختلفا نوعا ما، حيث يعتبر أنه إذا كانت القيم المستقبلية لـ Y_{1t} تسمح بشرح القيم الحالية لـ Y_{2t}، فإن Y_{2t} يسبب Y_{1t}. لدينا:

$$Y_{1t} = \phi_1^0 + \sum_{i=1}^{p}\phi_{1i}^1 Y_{1t-i} + \sum_{i=1}^{p}\phi_{1i}^2 Y_{2t-i} + \sum_{i=1}^{p}\beta_i^2 Y_{2t+i} + \varepsilon_{1t}$$

$$Y_{2t} = \phi_2^0 + \sum_{i=1}^{p}\phi_{2i}^1 Y_{1t-i} + \sum_{i=1}^{p}\phi_{2i}^2 Y_{2t-i} + \sum_{i=1}^{p}\beta_i^1 Y_{1t+i} + \varepsilon_{2t}$$

- Y_{2t} لا يسبب Y_{1t} إذا كانت الفرضية $H_0 : \beta_1^1 = \beta_2^1 = ...\beta_p^1 = 0$ مقبولة
- Y_{1t} لا يسبب Y_{2t} إذا كانت الفرضية $H_0 : \beta_1^2 = \beta_2^2 = ...\beta_p^2 = 0$ مقبولة

وهنا يتعلق الأمر باختبار فيشر الكلاسيكي (اختبار انعدام المعاملات).

<u>مثال 3:</u>

انطلاقا من النموذج VAR(1) المقدر في المثال السابق، نقوم بتطبيق اختباري Granger و Sims للسببية.

- اختبار Granger:

Y_{1t} لا تسبب $H_0\ Y_{2t}$:

نقدر النموذج UVAR و RVAR التاليين:

UVAR: $\hat{Y}_{1t} = 0.096 + 0.28Y_{1t-1} + 0.005Y_{2t-1}, R^2 = 0.99,$
$URSS = 1128.54$

RVAR: $\hat{Y}_{1t} = 12.57 + 0.58Y_{1t-1},\ R^2 = 0.69,\ RRSS = 1986.85$

نستعمل برمجية RATS 5.04 لاختبار فيشر (انعدام معامل Y_{2t-1}):

```
LINREG Y1
# CONSTANT Y1{1} Y2{1}
EXCLUDE
# Y2{1}
```

في هذه الحالة، نحسب إحصائية فيشر:

$$F^* = \frac{(RRSS - URSS)/c}{URSS/(T-k-1)} = \frac{(1986.85-1128)/1}{1128/(164-2-1)} = 122.58$$

نلاحظ أن $F^* > F_{0.05}$، أي نرفض فرضية العدم H_0، Y_{2t} تشرح معنويا المتغير Y_{1t} وهذا يعني أن هناك سببية وفق Granger.

بنفس الطريقة، نختبر السببية لـ Y_{1t} . نحو Y_{2t} على المعادلة 2 (أنظر المثال 2):

Y_{2t} لا تسبب Y_{1t} : H_0

أو بطريقة أخرى، نستعمل نسبة المعقولية لاختبار السببية وفق Granger ولهذا نحسب النسبة بالاستعانة بالبرنامج التالي:

```
SYSTEM(MODEL=UNRESTRICTED)
VARIABLES Y1 Y2 Y3
LAGS 1 TO 1
DET CONSTANT Y{1} Y{2}
END(SYSTEM)
ESTIMATE(RESIDS=UNRESIDS)
*
SYSTEM(MODEL=RESTRICTED)
VARIABLES Y1 Y2
LAGS 1 TO 1
DET CONSTANT Y1{1}
END(SYSTEM)
ESTIMATE(RESIDS=RESRESIDS)
*
RATIO (DEGREES=2)
# UNRESIDS
# RESRESIDS
```

أي:

$$L^* = (T-c)\times\left(\ln|\Sigma_{RVAR}| - \ln|\Sigma_{UVAR}|\right) = (164-1)(9.21-9.11) = 16.30 > \chi^2_{0.05}(2) = 5.99$$

- اختبار Sims:

نقدر النموذجين التاليين على نفس الفترة باستعمال التعليمة التالية:

```
LINREG Y1
#CONSTANT Y1{1} Y2{1} Y2{-1}
EXCLUDE
# Y2{-1}
```

H_0 : Y_{2t} لا تسبب Y_{1t}

UVAR: $\hat{Y}_{1t} = \text{-}1.025 + 0.25Y_{1t-1} - 0.18Y_{2t-1} + 0.29Y_{2t+1}$, $R^2 = 0.65$,
$URSS = 2428.65$

RVAR: $\hat{Y}_{1t} = 0.278 + 0.48Y_{1t-1} - 0.22Y_{2t-2}$, $R^2 = 0.43$, $RRSS = 3779.85$

$$F^* = \frac{(RRSS - URSS)/c}{URSS/(T-k-1)} = \frac{(3779.85-2428.65)/1}{2428.65/(164-3-1)} = 89.07$$

نلاحظ أن $F^* > F_{0.05}$، أي نرفض H_0. y_{2t+1} يشرح معنويا المتغير y_{1t}. هناك سببية وفق Sims.

نختبر أيضا السببية في الاتجاه المعاكس على المعادلة 2: H_0 : Y_{1t} لا تسبب Y_{2t}.

2.2. تحليل الصدمات ودوال الاستجابة **Impulse analysis**:

كما نعلم، نموذج VAR يُنمذج العلاقات الحركية بين مجموعة من المتغيرات المختارة لوصف ظاهرة اقتصادية خاصة. إن تحليل الصدمات ودوال الاستجابة يسمح بدراسة أثر صدمة معينة على متغيرات النظام. لنأخذ النموذج المقدر التالي:

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} \hat{\phi}_1^0 \\ \hat{\phi}_2^0 \end{pmatrix} + \begin{pmatrix} \hat{\phi}_{11}^1 & \hat{\phi}_{11}^2 \\ \hat{\phi}_{21}^1 & \hat{\phi}_{21}^2 \end{pmatrix} \begin{pmatrix} Y_{1t-1} \\ Y_{2t-1} \end{pmatrix} + \begin{pmatrix} \hat{\varepsilon}_{1t} \\ \hat{\varepsilon}_{2t} \end{pmatrix}$$

تغير في $\hat{\varepsilon}_{1t}$ خلال فترة زمنية معينة له نتيجة على Y_{1t} و Y_{1t+1} ثم على Y_{2t+1}، فإذا حدثت صدمة في اللحظة t على $\hat{\varepsilon}_{1t}$ تساوي 1 فإن أثرها يكون كالتالي:

في الفترة t: $\begin{pmatrix} \nabla Y_{1t} \\ \nabla Y_{2t} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

في الفترة $t+1$: $\begin{pmatrix} \nabla Y_{1t+1} \\ \nabla Y_{2t+1} \end{pmatrix} = \begin{pmatrix} \hat{\phi}_{11}^1 & \hat{\phi}_{11}^2 \\ \hat{\phi}_{21}^1 & \hat{\phi}_{21}^2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

في الفترة $t+2$: $\begin{pmatrix} \nabla Y_{1t+2} \\ \nabla Y_{2t+2} \end{pmatrix} = \begin{pmatrix} \hat{\phi}_{11}^1 & \hat{\phi}_{11}^2 \\ \hat{\phi}_{21}^1 & \hat{\phi}_{21}^2 \end{pmatrix} \begin{pmatrix} \nabla Y_{1t+1} \\ \nabla Y_{2t+1} \end{pmatrix}$

... الخ

تشكل هذه القيم المحسوبة دالة الاستجابة. تتميز طريقة دوال الاستجابة لحساب المضاعفات الديناميكية الموجودة بأنها تأخذ بعين الاعتبار مجموع العلاقات الديناميكية الموجودة، بحيث أنها تبين رد فعل نظام المتغيرات الداخلية على أثر حدوث صدمة في الأخطاء وحسب سيمس فإن دوال الاستجابة تبين أثر انخفاض وحيد ومفاجئ لمتغيرة على نفسها وعلى باقي متغيرات النظام في كل الأوقات. في هذه الحالة، نفترض أن البواقي مستقلة لكن هذه الفرضية نادرا ما تكون محققة، لأن في الواقع قد يوجد ارتباط بين الأخطاء العشوائية. إذا كان هناك ارتباط قوي بين صدمتين ε_{1t} و ε_{2t}، فإن صدمة ما على ε_{1t} حتما ستكون متبوعة بصدمة على ε_{2t}. في هذه الحالة معامل الارتباط سيؤكد على الصلة المشتركة بين البواقي $\hat{\varepsilon}_{1t}$ و $\hat{\varepsilon}_{2t}$ ولكن لا تشير إلى اتجاه السببية.

إن مشكل الارتباط المشترك للأخطاء وأثر صدمة على متغير ما قد يعالج بالبحث عن تمثيل الأخطاء بصفة شاقولية Orthogonal (مستقلة فيما بينها). لنعتبر تقسيم Σ :

$$\Sigma = PP'$$

يتعلق المر هنا بتقسيم Choleski حيث P تعبر عن مصفوفة مثلثية في الأعلى مع عناصرها القطرية موجبة. يمكن كتابة الصيغة $VMA(\infty)$ على الشكل التالي:

$$Y_t = \mu + \sum_{i=0}^{\infty} C_i P P^{-1} \varepsilon_{t-i} = \mu + \sum_{i=0}^{\infty} M_i \nu_{t-i}$$

مع: $M_i = C_i P$ و $\nu_t = P^{-1}\varepsilon_t$.

من السهل التأكد من أن للأخطاء ν_t مصفوفة تباين-تباين مشترك تساوي المصفوفة الأحادية. أعمدة M_i تمثل استجابة النظام بالنسبة لصدمة مستقلة و طبيعية على خطأ متغير ما بعد t فترة زمنية[1].

وكمثال على ذلك، نأخذ نموذج VAR لمتغيرين:

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} \phi_{11}^1 & \phi_{11}^2 \\ \phi_{21}^1 & \phi_{21}^2 \end{pmatrix} \begin{pmatrix} Y_{1t-1} \\ Y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

مع: $\mathrm{var}(\varepsilon_{1t}) = \sigma_{\varepsilon}^2 \ , \ \mathrm{var}(\varepsilon_{2t}) = \sigma_{\varepsilon_2}^2$ و $\mathrm{cov}(\varepsilon_{1t}, \varepsilon_{2t}) = k \neq 0$.

بحساب $Y_{1t} - \left(\sigma_{\varepsilon_1}^2 / k\right) Y_{2t}$ نحصل على:

$$Y_{2t} = \left(k / \sigma_{\varepsilon_1}^2\right) Y_{1t} + \left(\phi_{21}^1 - \phi_{11}^1 \times k / \sigma_{\varepsilon_1}^2\right) Y_{1t-1} + \left(\phi_{21}^2 - \phi_{11}^2 \times k / \sigma_{\varepsilon_1}^2\right) Y_{2t-1} + \varepsilon_{2t} - \left(k / \sigma_{\varepsilon_1}^2\right) \varepsilon_{1t}$$

نضع: $\nu_t = \varepsilon_{2t} - \left(k / \sigma_{\varepsilon_1}^2\right) \varepsilon_{1t}$. لدينا:

$$\mathrm{cov}(\varepsilon_{1t}, \nu_t) = E(\varepsilon_{1t} \nu_t) = \mathrm{cov}(\varepsilon_{1t}, \varepsilon_{2t}) - k / \sigma_{\varepsilon_1}^2 E(\varepsilon_{1t}^2) = k - k = 0$$

أصبحت الأخطاء غير مرتبطة (شاقولية) ويمكن تحليل الصدمات على المعادلتين التاليتين:

1- أنظر Sims (1981).

$$Y_{1t} = \phi_{11}^{1} Y_{1t-1} + \phi_{11}^{2} Y_{2t-1} + \varepsilon_{1t}$$
$$Y_{2t} = \left(k/\sigma_{\varepsilon_1}^{2}\right) Y_{1t} + \left(\phi_{21}^{1} - \phi_{11}^{1} \times k/\sigma_{\varepsilon_1}^{2}\right) Y_{1t-1} + \left(\phi_{21}^{2} - \phi_{11}^{2} \times k/\sigma_{\varepsilon_1}^{2}\right) Y_{2t-1} + \varepsilon_{2t} - \left(k/\sigma_{\varepsilon_1}^{2}\right) \varepsilon_{1t}$$

3.2. تحليل التباين Variance Decomposition:

يهدف تحليل تباين خطأ التنبؤ إلى حساب وتحديد مدى مساهمتها في تباين الخطأ. رياضيا، نستطيع كتابة تباين خطأ التنبؤ لفترة معينة h بدلالة تباين الخطأ الخاص بكل متغير على حدا. ولمعرفة وزن أو نسبة مشاركة كل تباين نقوم بقسمة هذا التباين على تباين خطأ التنبؤ الكلي.

بعدما تصبح الصدمات طبيعية و شاقولية، يتم تحليل الاستجابة بواسطة النموذج:

$$Y_t = \mu + \sum_{i=0}^{\infty} M_i v_{t-i}$$

خطأ التنبؤ في الأفق h يعطى بالعلاقة التالية:

$$Y_{t+h} - E_t(Y_{t+h}) = \sum_{i=0}^{h-1} M_i v_{t+h-i}$$

نقوم بتقسيم خطأ التنبؤ من أجل كل مركبة لـ Y_t . التي نرمز إليها بـ Y_{jt} . لدينا:

$$Y_{j,t+h} - E_t(Y_{j,t+h}) = \sum_{i=0}^{h-1} \left(m_{j1,i} v_{1,t+h-i} + m_{j2,i} v_{2,t+h-i} + \ldots\ldots + m_{jm,i} v_{m,t+h-i} \right)$$

حيث $m_{j1,i}$ يعبر عن العنصر $(j,1)$ الخاص بالمصفوفة M_j . يمكن التعبير عنه بطريقة مختلفة:

$$Y_{j,t+h} - E_t(Y_{j,t+h}) = \sum_{k=1}^{n} \left(m_{jk,1} v_{k,t+h} + \ldots\ldots + m_{jk,h-1} v_{k,t+1} \right)$$

بما أن الأخطاء v لا تشكل أي ارتباط وذات تباين يساوي 1، يسهل علينا حساب تباين خطأ التنبؤ:

$$E\left(Y_{j,t+h} - E_t(Y_{j,t+h})\right)^2 = \sum_{k=1}^{n} \left(m_{jk,1}^2 + \ldots\ldots + m_{jk,h-1}^2 \right)$$

مع:

$$m_{jk,1}^2 + \ldots\ldots + m_{jk,h-1}^2 = \sum_{i=0}^{h-1} \left(e_j' M_i e_k \right)^2$$

حيث e_i يمثل العمود رقم i للمصفوفة الأحادية والتي تعبر عن مساهمة خطأ المتغير k في تباين خطأ التنبؤ في الأفق h للمتغير j. للحصول على التحليل The decomposition بالنسبة المئوية، نجعل العبارة على الشكل:

$$\sum_{k=1}^{n}\sum_{i=0}^{h-1} m_{jk,i}^2$$

لنأخذ ثانية النموذج VAR(1) لمتغيرين، تباين خطأ التنبؤ لـ . Y_{1t+h} يكتب كما يلي:

$$\sigma_{Y_1}^2(h) = \sigma_{Y_1}^2\left[m_{11}^2(0) + m_{11}^2(1) + ...+ m_{11}^2(h-1)\right] + \sigma_{Y_2}^2\left[m_{22}^2(0) + m_{22}^2(1) + ...+ m_{22}^2(h-1)\right]$$

حيث m_{ii} هي عناصر المصفوفة M.

في الأفق h، تحليل تباين الأخطاء (Y_{1t} على Y_{2t})، بالنسبة المئوية (على الترتيب Y_{2t} على Y_{1t}) يعطى بالصيغة:

$$\frac{\sigma_{Y_1}^2\left[m_{11}^2(0) + m_{11}^2(1) + ...+ m_{11}^2(h-1)\right]}{\sigma_{Y_1}^2(h)}$$

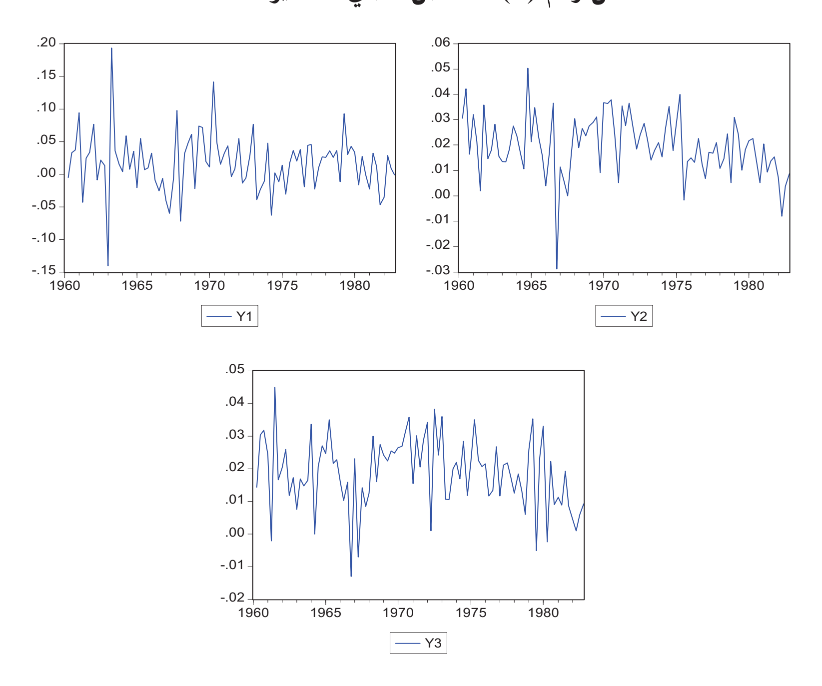
$$\frac{\sigma_{Y_2}^2\left[m_{22}^2(0) + m_{22}^2(1) + ...+ m_{22}^2(h-1)\right]}{\sigma_{Y_1}^2(h)}$$

إذا كانت صدمة معينة على ε_{1t} لا تؤثر على تباين خطأ Y_{1t} مهما تكن h، فمن المحتمل أن Y_{2t} متغير خارجي باعتبار أن Y_{2t} و ε_{1t} مستقلان. وعكس ذلك، إذا كانت لصدمة معينة على ε_{1t} أثر كبير على تباين خطأ Y_{2t} فإن هذا الأخير متغير داخلي.

مثال 4:

لنعتبر نموذج VAR(2) وانطلاقا من معطيات اقتصادية تظهر في الشكل (1) Y_{1t}, Y_{2t}, Y_{3t} نقوم بحساب دوال الاستجابة وتحليل التباين. النتائج معطاة بالاستعانة ببرمجية Eviews 5.0 ونشير هنا إلى أن RATS 5.04 يتضمن برنامجا كاملا IMPULSES.PRG يسمح بحساب العناصر المطلوبة.

الشكل رقم (1): التمثيل البياني للمتغيرات الثلاثة



الجدول (1): دالة الاستجابة

Response of Y1:

Period	Y1	Y2	Y3
1	0.043880	0.000000	0.000000
	(0.00363)	(0.00000)	(0.00000)
2	-0.011369	0.006122	0.006944
	(0.00521)	(0.00501)	(0.00460)
3	-0.000941	0.004840	0.003002
	(0.00521)	(0.00515)	(0.00456)
4	0.004708	0.001983	-0.002947
	(0.00330)	(0.00305)	(0.00330)
5	0.001319	0.001425	0.001887
	(0.00191)	(0.00207)	(0.00171)
6	-0.000793	-0.000419	0.000482
	(0.00140)	(0.00168)	(0.00106)
7	0.000518	0.001065	-0.000349
	(0.00067)	(0.00101)	(0.00067)
8	0.000258	9.87E-05	0.000271
	(0.00047)	(0.00040)	(0.00032)
9	3.60E-05	-3.15E-05	1.24E-05
	(0.00034)	(0.00035)	(0.00026)
10	2.11E-05	0.000173	1.62E-05
	(0.00012)	(0.00019)	(0.00013)

Response of Y2:

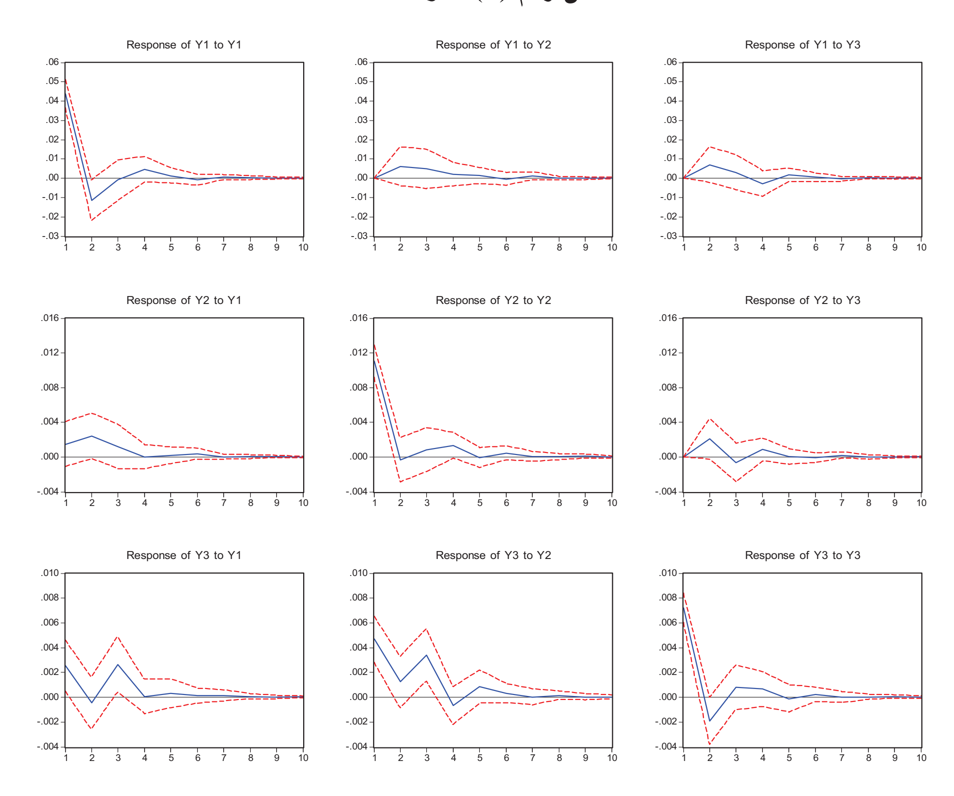
Period	Y1	Y2	Y3
1	0.001476	0.011045	0.000000
	(0.00130)	(0.00091)	(0.00000)
2	0.002435	-0.000333	0.002084
	(0.00131)	(0.00127)	(0.00117)
3	0.001198	0.000843	-0.000637
	(0.00129)	(0.00127)	(0.00112)
4	2.15E-05	0.001351	0.000865
	(0.00069)	(0.00074)	(0.00065)
5	0.000170	-8.53E-05	5.79E-05
	(0.00048)	(0.00058)	(0.00044)
6	0.000357	0.000456	-9.63E-05
	(0.00032)	(0.00040)	(0.00027)
7	1.09E-05	6.28E-05	0.000199
	(0.00014)	(0.00027)	(0.00018)
8	3.01E-05	2.76E-05	-2.49E-05
	(0.00011)	(0.00017)	(0.00013)
9	4.84E-05	9.68E-05	8.35E-06
	(6.1E-05)	(0.00010)	(6.8E-05)
10	9.58E-06	-4.94E-06	2.56E-05
	(3.9E-05)	(7.3E-05)	(4.6E-05)

Response of Y3:

Period	Y1	Y2	Y3
1	0.002539	0.004692	0.007224
	(0.00103)	(0.00093)	(0.00060)
2	-0.000445	0.001245	-0.001907
	(0.00106)	(0.00104)	(0.00095)
3	0.002646	0.003397	0.000795
	(0.00113)	(0.00106)	(0.00090)
4	6.20E-05	-0.000658	0.000657
	(0.00071)	(0.00076)	(0.00071)
5	0.000312	0.000860	-0.000114
	(0.00058)	(0.00067)	(0.00055)
6	0.000119	0.000312	0.000231
	(0.00030)	(0.00038)	(0.00028)
7	0.000149	2.00E-05	3.17E-06
	(0.00022)	(0.00032)	(0.00022)
8	5.91E-05	0.000147	2.18E-05
	(0.00010)	(0.00017)	(0.00010)
9	8.66E-06	2.51E-05	4.68E-05
	(7.0E-05)	(0.00013)	(6.6E-05)
10	2.66E-05	2.51E-05	-1.07E-05
	(4.0E-05)	(6.8E-05)	(4.7E-05)

Cholesky Ordering: Y1 Y2 Y3
Standard Errors: Analytic

الشكل رقم (2): دوال الاستجابة



الجدول (2): تحليل التباين

Variance Decomposition of Y1:

Period	S.E.	Y1	Y2	Y3
1	0.043880	100.0000	0.000000	0.000000
2	0.046264	95.99597	1.751092	2.252935
3	0.046623	94.56487	2.802134	2.632994
4	0.046995	94.07918	2.936115	2.984706
5	0.047072	93.84637	3.018058	3.135568
6	0.047083	93.83084	3.024575	3.144584
7	0.047100	93.77844	3.073651	3.147905
8	0.047101	93.77511	3.073882	3.151006
9	0.047101	93.77507	3.073924	3.151009
10	0.047102	93.77379	3.075228	3.150978

Variance Decomposition of Y2:

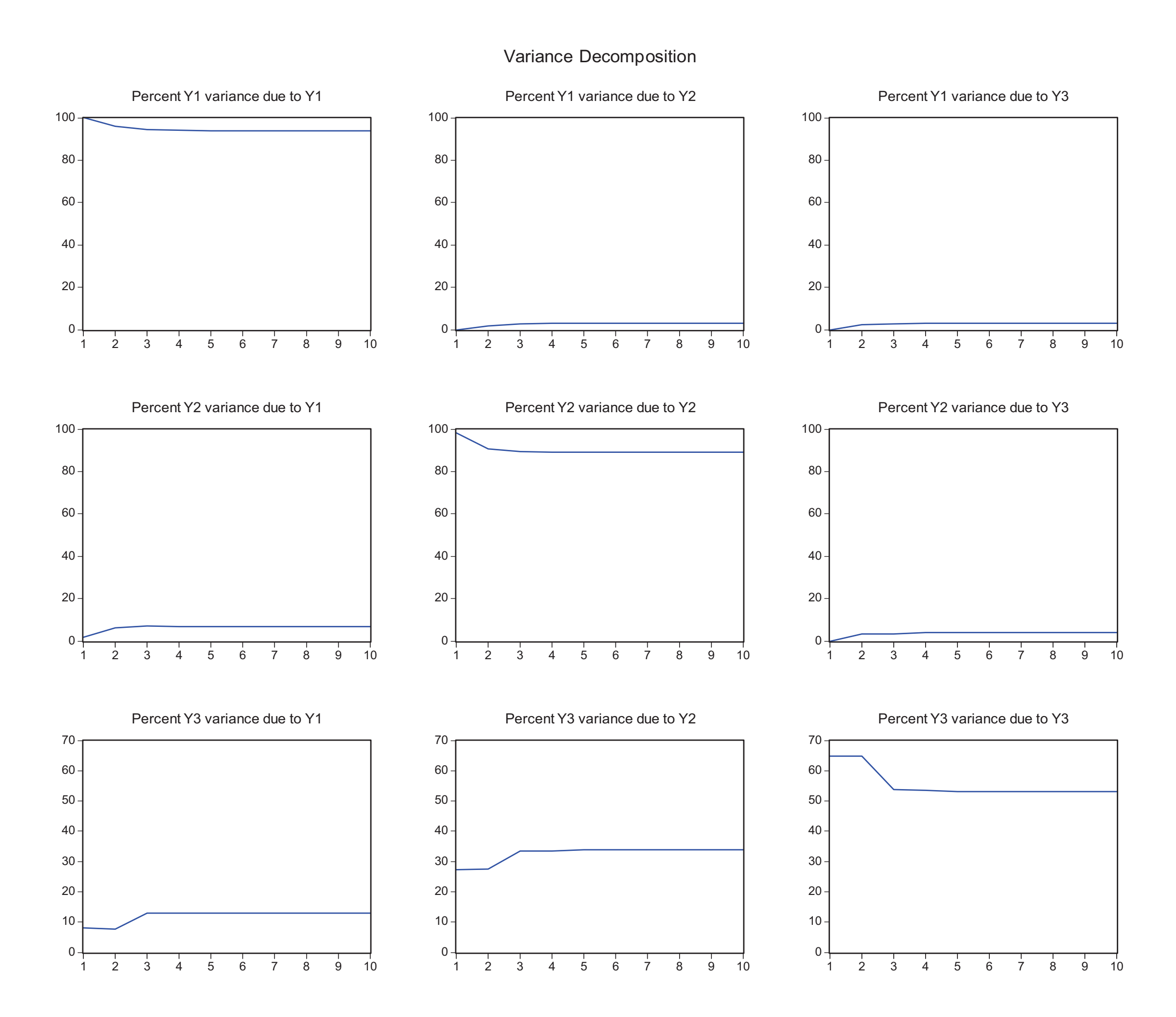
Period	S.E.	Y1	Y2	Y3
1	0.011143	1.753616	98.24638	0.000000
2	0.011600	6.024526	90.74698	3.228490
3	0.011709	6.959247	89.57624	3.464512
4	0.011819	6.831280	89.23210	3.936621
5	0.011820	6.850095	89.21200	3.937900
6	0.011835	6.924321	89.14084	3.934843
7	0.011837	6.922243	89.11580	3.961954
8	0.011837	6.922776	89.11489	3.962333
9	0.011837	6.923867	89.11408	3.962050
10	0.011837	6.923895	89.11361	3.962497

Variance Decomposition of Y3:

Period	S.E.	Y1	Y2	Y3
1	0.008981	7.995029	27.29210	64.71288
2	0.009275	7.724763	27.38483	64.89040
3	0.010257	12.97288	33.36411	53.66301
4	0.010299	12.87033	33.49875	53.63092
5	0.010341	12.85881	33.92442	53.21677
6	0.010349	12.85218	33.96299	53.18483
7	0.010350	12.87021	33.95619	53.17360
8	0.010351	12.87041	33.96822	53.16138
9	0.010351	12.87013	33.96789	53.16198
10	0.010351	12.87062	33.96801	53.16137

Cholesky Ordering: Y1 Y2 Y3

الشكل رقم (3): تحليل التباين



نعتبر أن خطأ ما على Y_{1t} لا يؤثر بشكل مباشر وآني على Y_{2t} و Y_{3t}، لكن خطأ Y_{2t} له أثر على Y_{1t} ونفس الشيء بالنسبة لـ Y_{3t}. من خلال الشكل (2)، نلاحظ مثلا أن مدى الصدمة على Y_{1t} تساوي مرة الخطأ المعياري، الصدمة على Y_{1t} تساوي 0.043 بعد مرور فترة زمنية واحدة. هذه الصدمة تنعكس على السيرورات الثلاثة بمسار ضعيف

(بطيء). الصدمة على Y_{2t} تؤثر على Y_{1t} بصفة آنية وتنعكس بعد ذلك على السيرورات الثلاثة. كما أن الصدمة على Y_{3t} تؤثر هي الأخرى على مسار Y_{2t} بطريقة آنية.

تحليل التباين Decomposition Variance المبين في الجدول (2) (أنظر أيضا الشكل (3)) يشير إلى أن تباين خطأ التنبؤ الخاص بـ . Y_{1t} يعادل تقريبا 94% من أخطاء هذا المتغير Innovations و 2% من تلك المتعلقة بـ . Y_{2t} و 3% من تلك الخاصة بـ . Y_{3t}. تباين خطأ التنبؤ لـ . Y_{2t} يساهم بنسبة 90% من أخطاء Y_{2t} و 6% من تلك الخاصة بـ . Y_{1t} و 3% من تلك المتعلقة بـ . Y_{3t}. أما تباين خطأ التنبؤ لـ . Y_{3t} يساوي 54% من أخطائه و 32% من تلك الخاصة بـ . Y_{2t} و 11% من أخطاء Y_{1t}.

3. التكامل المشترك ونموذج تصحيح الخطأ Cointegration and VECM

1.3. مفهوم التكامل المشترك Concept of Cointegration

ظهرت تقنية التكامل المشترك في أواسط الثمانينات على يد (1983) Granger و (1987) Engle، وارتكز تطورها قبل كل شيء على صحة فرضية استقرارية السلاسل الزمنية، وهي ناتجة عن عملية دمج بين تقنية بوكس-جينكيتر والتقارب الحركي (الديناميكي) لنماذج تصحيح الخطأ. ترتكز هذه التقنية على السلاسل الزمنية غير المستقرة، في حين تكون التركيبات الخطية التي فيما بينها مستقرة، وجود التكامل المشترك مرتبط باختبارات الجذر الوحدوي للتحقق من استقرار السلاسل، كما تسمح هذه الاختبارات من التأكد من وجود تكامل مشترك أي التقارب بين سيرورات السلاسل الزمنية.

1.1.3. خصائص درجة تكامل سلسلة زمنية وشروط التكامل المشترك:

كما رأينا في الفصل السابق، تكون سلسلة زمنية معينة متكاملة من الدرجة d $X_t \rightarrow I(d)$، إذا تم حساب الفروقات d مرة من أجل جعل السلسلة مستقرة.

لتكن سلسلة زمنية X_{1t} مستقرة و سلسلة أخرى X_{2t} متكاملة من الدرجة 1:

$$X_{1t} \rightarrow I(0)$$
$$X_{2t} \rightarrow I(1)$$

تعتبر السلسلة $Y_t = X_{1t} + X_{2t}$ غير مستقرة لأننا قمنا بجمع سلسلتين الأولى مستقرة و الثانية غير مستقرة تحتوي على اتجاه عام.

إذا كانت لدينا سلسلتان X_{1t} و X_{2t} متكاملتان من الدرجة d، فما هي إذن درجة تكامل $X_{1t} + X_{2t}$ ؟ وما هي درجة تكامل $\alpha X_{1t} + \beta X_{2t}$ ؟ فالنتيجة تتوقف على معرفة إشارات المعاملين α و β وترتبط بوجود ديناميكية غير مستقرة مشتركة.

نقول أن هناك تكامل مشترك بين السلسلتين X_t و Y_t " The two series are cointegrated" إذا تضمنتا اتجاها عاما عشوائيا بنفس درجة التكامل d و توليفة خطية للسلسلتين تسمح بالحصول على سلسلة ذات درجة تكامل أقل.

ليكن:

$$X_t \to I(d)$$
$$Y_t \to I(d)$$

بحيث $\alpha_1 X_t + \alpha_2 Y_t \to I(d-b)$ مع $d \geq b \geq 0$. نرمز بـ: $X_t, Y_t \to CI(d,b)$
حيث $[\alpha_1 \quad \alpha_2]$ يسمى بشعاع التكامل المشترك.

في الحالة العامة، إذا كان لدينا k متغير، فإن:

$$X_{1t} \to I(d)$$
$$X_{2t} \to I(d)$$
$$....$$
$$X_{kt} \to I(d)$$

نضع: $X_t = [X_{1t} \quad X_{2t} \quad \cdots \quad X_{kt}]$

إذا وجد شعاع تكامل مشترك $\alpha = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_k]$ ذو بعد $(k,1)$ بحيث $\alpha X_t \to I(d-b)$، فإن المتغيرات التي عددها k تحمل خاصية التكامل المشترك وشعاع التكامل المشترك هو α. نضع: $X_t \to CI(d,b)$ مع $b > 0$.

2.1.3. نموذج تصحيح الخطأ Error Correction Model (ECM):

ندرس الحالة التي يكون بين X_t و Y_t تكامل مشترك حيث $X_t, Y_t \rightarrow CI(1,1)$ و $[\beta,-1]$. نضع $\beta = -\frac{\alpha_1}{\alpha_2}$ شعاع التكامل المشترك أي $\beta X_t - Y_t \rightarrow I(0)$. في هذا النوع من النمذجة، كون السلسلتين متكاملتين cointegrated وغير مستقرتين يزيد من مشكل التقدير. المعنوية الإحصائية للنموذج هي السبب في كون أن السلسلتين غير مستقرتين (أي هناك تكامل مشترك). إن استعمال الانحدار المباشر لـ Y_t على X_t يعتبر غير ممكن باعتبار أن العلاقة المفترضة انطلاقا من هذا الانحدار ليست واقعية، فينجم عن ذلك الحصول على علاقة بين اتجاهين Two trends.

يكمن المشكل الأساسي في سحب العلاقة المشتركة للتكامل المشترك (الاتجاه العام المشترك) من جهة ومن جهة أخرى البحث عن الارتباط الحقيقي بين المتغيرين وهو هدف نموذج تصحيح الخطأ ECM، فهو يجمع بين النموذج الساكن $\beta_1 \nabla X_t$ و النموذج الديناميكي (الحركي) $\beta_2 (Y_{t-1} - \beta X_{t-1})$. ليكن:

$$\nabla Y_t = \beta_1 \nabla X_t + \beta_2 (Y_{t-1} - \beta X_{t-1})$$
$$I(0) \qquad I(0) \qquad\qquad I(0)$$

إضافة إلى العلاقة طويلة المدى، يسمح نموذج تصحيح الخطأ في دمج التقلبات قصيرة المدى. المعامل β_2 الذي ينبغي أن يكون سالبا يمثل قوة الجذب (الرجوع) نحو التوازن طويل المدى.

2.3. اختبار التكامل المشترك وتقدير نموذج تصحيح الخطأ

يرتكز اختبار التكامل المشترك على الخوارزمية التي اقترحها Engle and Granger (1987) وهي طريقة على مرحلتين:

الخطوة الأولى: اختبار درجة تكامل المتغيرين:

الشرط الضروري للتكامل يتمثل في أن السلسلتين ينبغي أن تكونا متكاملتين من نفس الدرجة (الرتبة). إذا كانتا غير متكاملتين من نفس الدرجة (الرتبة)، فهذا يعني أنهما لا تحققان خاصية التكامل المشترك Two series are not cointegrated.

لا بد من تحديد نوع الاتجاه العام بعناية (ثابت أو عشوائي) لكل متغير ثم درجة التكامل d للسلسلتين المدروستين. إذا كانت السلسلتان متكاملتين من نفس الدرجة (الرتبة)، فهناك تكامل مشترك بينهماThe two series are cointegrated.

الخطوة الثانية: تقدير العلاقة طويلة المدى:

إذا كان الشرط الضروري محققا، فينبغي تقدير العلاقة طويلة المدى بين المتغيرين $Y_t = a_1 X_t + a_0 + \varepsilon_t$ بطريقة المربعات الصغرى العادية OLS.

من أجل قبول علاقة التكامل المشترك، يجب أن تكون سلسلة بواقي التقدير $\hat{\varepsilon}_t$ مستقرة (اختبار الاستقرارية يتم عن طريق اختبار Dickey-Fuller أو Philips-Perron أو تمثيل دالة الارتباط الذاتي للبواقي) حيث $\hat{\varepsilon}_t = Y_t - \hat{a}_1 X_t - \hat{a}_0$.

في هذه الحالة، لا يمكن استخدام جداول Dickey-Fuller، فالاختبار يتم على البواقي انطلاقا من العلاقة الساكنة وليس على البواقي الحقيقية من علاقة التكامل المشترك. قام Mackinnon (1991) بمحاكاة الجداول التي تعتمد على عدد المشاهدات وعدد المتغيرات المستقلة التي تظهر في العلاقة الساكنة.

إذا كانت البواقي مستقرة، يمكن إذن تقدير نموذج تصحيح الخطأ. عندما تكون السلسلتان غير مستقرتين وتحمل صفة التكامل المشترك، فلا بد من تقدير العلاقة انطلاقا من نموذج ECM.

لتكن السلسلتان X_t و Y_t متكاملتين من الدرجة الأولى أي $X_t \to I(1)$ و $Y_t \to I(1)$، تقدير العلاقة طويلة المدى تشير إلى استقرارية البواقي، نضع $X_t, Y_t \to CI(1,1)$. نقدر نموذج ECM بإتباع الخطوات التالية:

الخطوة الأولى: تقدير العلاقة طويلة المدى بطريقة OLS:

$$Y_t = \hat{\alpha} + \hat{\beta} X_t + \hat{\varepsilon}_t$$

الخطوة الثانية: تقدير العلاقة قصيرة المدى (النموذج الديناميكي) بطريقة OLS:

$$\nabla Y_t = \alpha_1 \nabla X_t + \alpha_2 \hat{\varepsilon}_{t-1} + u_t \quad , \quad \alpha_2 < 0$$

يجب أن يكون المعامل α_2 معنويا سالبا، إذا لم يكن كذلك، يجب رفض نمذجة ECM. ميكانيزم تصحيح الخطأ الذي يؤدي إلى العلاقة طويلة المدى يذهب في الاتجاه المعاكس ويبتعد عن الهدف طويل المدى. تعطي التقنية على مرحلتين مقدرا متقاربا Convergent لمعاملات النموذج والأخطاء المعيارية للمعاملات تفسر بطريقة كلاسيكية. يمكن القول أيضا أن نموذج ECM لا يطرح أي مشكل خاص ويحتاج فقط إلى طريقة OLS.

مثال 5:

لتكن المعطيات الإحصائية X_t و Y_t المبينة في الجدول (3) أدناه والمطلوب تقدير العلاقة بين هذين المتغيرين واختبار التكامل المشترك وفي هذه الحالة تقدير نموذج تصحيح الخطأ.

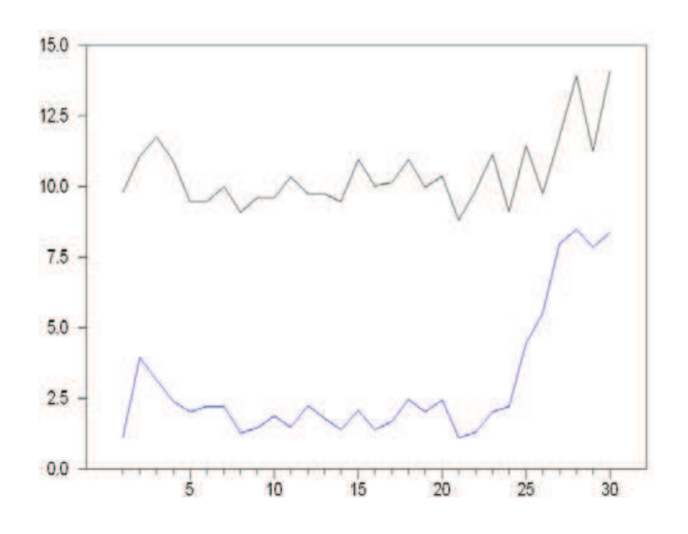
الجدول (3): القيم المشاهدة لكل من X_t و Y_t

المشاهدات	X_t	Y_t
1	01.11	09.79
2	03.96	11.08
3	03.16	11.75
4	02.41	10.87
5	02.02	09.46
6	02.23	09.46
7	02.23	10.00
8	01.30	09.10
9	01.45	09.59
10	01.90	09.59
11	01.49	10.34
12	02.25	09.73
13	01.81	09.74

Y_t	X_t	المشاهدات
09.46	01.39	14
10.95	02.09	15
10.03	01.39	16
10.13	01.69	17
10.95	02.48	18
09.98	02.04	19
10.36	02.46	20
08.80	01.11	21
09.88	01.32	22
11.13	02.04	23
09.12	02.23	24
11.45	04.46	25
09.74	05.51	26
11.73	07.95	27
13.92	08.51	28
11.24	07.85	29
14.09	08.38	30

نمثل بيانيا هاتين السلسلتين:

الشكل رقم (4): التمثيل البياني للسلسلتين



من خلال الشكل (4)، نلاحظ أن السلسلتين غير مستقرتين، فلا بد إذن التأكد من ذلك باختبار الجذر الوحدوي (Dickey-Fuller، Philips-Perron و KPSS). الجدول التالي يعطي نتائج هذه الاختبارات:

الجدول (4): نتائج اختبارات الجذر الوحدوي

	نوع الاختبار	**نوع النموذج**	**القيمة المحسوبة (X_t) (القيمة الحرجة)**	**القيمة المحسوبة (Y_t) (القيمة الحرجة)**
السلسلة الأصلية	اختبار ADF جذر وحدوي: H_0	(1)	1.056 (-1.952)	0.317 (-1.952)
		(2)	0.059 (-2.967)	-2.875 (-2.967)
		(3)	-0.839 (-3.574)	-3.434 (-3.574)
	اختبار Philips-Perron (PP) جذر وحدوي: H_0	(1)	1.076 (-1.952)	0.851 (-1.952)
		(2)	-0.188 (-2.967)	-3.271 (-2.967)
		(3)	-0.879 (-3.574)	-3.682 (-3.574)

السلسلة المحولة (الفروقات من الدرجة الأولى)	اختبار ADF جذر وحدوي: H_0	(1)	5.694- (-1.953)	-9.450 (-1.953)
		(2)	-5.714 (-2.971)	-9.363 (-2.971)
	اختبار Philips-Perron (PP) جذر وحدوي: H_0	(1)	-5.569 (-1.953)	-10.016 (-1.953)
		(2)	-5.599 (-2.971)	-9.363 (-2.971)

من خلال نتائج الاختبار، يتضح أن السلسلتين غير مستقرتين تحتويان على جذر وحدوي باعتبار أن القيم المحسوبة أقل تماما من القيم الحرجة لـ Mackinnon، ففي هذه الحالة نرفض أيضا فرضية الاتجاه العام التحديدي TS. أما السلسلتان المحولتان عن طريق

الفروقات من الدرجة الأولى مستقرتان أي أن X_t و Y_t متكاملتان integrated *I(1)*، فهذا يعني أن هناك احتمال وجود تكامل مشترك. نقدر العلاقة طويلة المدى $Y_t = \alpha + \beta X_t + \varepsilon_t$. نتائج التقدير تظهر باستخدام RATS 5.04:

```
linreg y / resids
#constant x

Linear Regression - Estimation by Least Squares
Dependent Variable Y
Usable Observations     30     Degrees of Freedom    28
Centered R**2      0.625319      R Bar **2   0.611938
Uncentered R**2    0.994864      T x R**2      29.846
Mean of Dependent Variable      10.448666667
Std Error of Dependent Variable  1.252880084
Standard Error of Estimate       0.780477128
Sum of Squared Residuals        17.056047331
Regression F(1,28)                   46.7303
Significance Level of F           0.00000020
Durbin-Watson Statistic             2.290779

   Variable                      Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                  9.1414468012 0.2384798068     38.33216  0.00000000
2.  X                         0.4346774104 0.0635869286      6.83596  0.00000020
```

يتم اختبار التكامل المشترك انطلاقا من بواقي التقدير $\hat{\varepsilon}_t$. علينا أن نتأكد من أن البواقي مستقرة باستخدام إحصائيتي Dickey-Fuller و Philips-Perron:

الجدول (4): نتائج اختبارات الجذر الوحدوي على البواقي

اختبار Philips-Perron		اختبار Dickey-Fuller	
النموذج (2)	النموذج (1)	النموذج (2)	النموذج (1)
-6.138 (2.967)	-6.254 (-1.952)	-6.138 (-2.967)	-6.254 (-1.952)

من الملاحظ أن سلسلة البواقي مستقرة وهذا ما يقودنا إلى تقدير نموذج تصحيح الخطأ ECM بعد حساب البواقي في الفترة السابقة $\hat{\varepsilon}_{t-1} = Y_{t-1} - 0.434X_{t-1} - 9.141$. نتائج تقدير العلاقة قصيرة المدى $\nabla Y_t = \alpha_1 \nabla X_t + \alpha_2 \hat{\varepsilon}_{t-1} + u_t$ مبينة باستعمال RATS 5.04:

```
diff y / ydiff
diff x / xdiff

linreg ydiff / res
#xdiff resids{1}

Linear Regression - Estimation by Least Squares
Dependent Variable YDIFF
Usable Observations     29     Degrees of Freedom    27
Centered R**2     0.693325      R Bar **2   0.681967
Uncentered R**2   0.696968      T x R**2      20.212
Mean of Dependent Variable      0.1482758621
Std Error of Dependent Variable 1.3763774113
Standard Error of Estimate      0.7762000419
Sum of Squared Residuals        16.267135634
Durbin-Watson Statistic             1.898350

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  XDIFF                    0.397977883  0.159600286     2.49359  0.01907071
2.  RESIDS{1}               -1.240196814  0.220389841    -5.62729  0.00000568
```

من الملاحظ أن معامل $\hat{\varepsilon}_{t-1}$ معنويا سالب وهذا يعني أنه يجب قبول نموذج تصحيح الخطأ أي وجود تكامل مشترك بين X_t و Y_t.

3.3. تعميم التكامل المشترك و نموذج تصحيح الخطأ المتعدد VECM:

1.3.3. التكامل المشترك بين k متغير وتقدير نموذج تصحيح الخطأ:

ليكن النموذج القياسي لـ k متغير مفسر (مستقل):

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + ... + \beta_k X_{tk} + \varepsilon_t$$

إذا كانت المتغيرات Y_t و X_{tj} غير مستقرة حيث $j = 1,2,...,k$ (متكاملة من الدرجة الأولى مثلا، فهناك احتمال وجود تكامل مشترك. فوجود تكامل مشترك محتمل يعني أن

المتغيرات يجب أن تكون غير مستقرة ولها نفس درجة التكامل. التقدير بطريقة OLS يسمح إذن بحساب بواقي التقدير:

$$\hat{\varepsilon}_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_{t1} - \hat{\beta}_2 X_{t2} - ... - \hat{\beta}_k X_{tk}$$

إذا كانت سلسلة بواقي التقدير مستقرة، فإننا نقبل فرضية التكامل المشترك بين المتغيرات. اختبارات الاستقرارية لـ Dickey-Fuller . يجب أن تتم انطلاقا من القيم الحرجة المستعملة من طرف Mackinnon (1991) بدلالة عدد المتغيرات الكلية للنموذج. يعطى شعاع التكامل المشترك بـ : $\begin{bmatrix}1 & -\hat{\beta}_0 & -\hat{\beta}_1 & \cdots & -\hat{\beta}_k\end{bmatrix}$.

مع ذلك، الحالة التي يكون فيها نموذج متعدد المتغيرات أكثر تعقيدا من الحالة التي يكون فيها إلا متغيرين وذلك بسبب الاختيار عن طريق التوفيقات لشعاع التكامل المشترك وهذا يعني أنه مثلا إذا كانت المتغيرات Y_t, X_{t1}, X_{t2} و X_{t3} متكاملة بالزوج Cointegrated $Y_t, X_{t1} \rightarrow CI(1,1)$ و $X_{t2}, X_{t3} \rightarrow CI(1,1)$، فإن التوليفة الخطية لهذه المتغيرات هي $I(0)$. لدينا: $\hat{\varepsilon}_t^1 = Y_t - \hat{\alpha}_0 - \hat{\alpha}_1 X_{t1}$ و $\hat{\varepsilon}_t^2 = X_{t2} - \hat{\gamma}_0 - \hat{\gamma}_1 X_{t3}$ وكنتيجة لذلك:

$$\hat{\varepsilon}_t = \hat{\varepsilon}_t^1 + \hat{\varepsilon}_t^2 = Y_t - \hat{\alpha}_0 - \hat{\alpha}_1 X_{t1} + X_{t2} - \hat{\gamma}_0 - \hat{\gamma}_1 X_{t3} \rightarrow I(0)$$

في هذه الحالة نحصل على شعاع آخر للتكامل المشترك $\begin{bmatrix}1 & -\hat{\alpha}_0 & -\hat{\gamma}_0 & -\hat{\alpha}_1 & 1 & -\hat{\gamma}_1\end{bmatrix}$. في الحالة العامة، في نموذج ذي k متغير مستقل ومتغير تابع (أي $k+1$ متغير)، قد يوجد k شعاع تكامل مشترك مستقل خطيا. عدد أشعة التكامل المشترك المستقلة خطيا يسمى بدرجة التكامل المشترك.

إذا كانت للمتغيرات نفس درجة التكامل، فوجود شعاع تكامل مشترك واحد ممكن جدا وإذا لم تكن كلها لها نفس درجة التكامل، فمن المؤكد أنه لا يوجد شعاع تكامل مشترك وحيد.

عند اختبار التكامل المشترك هناك حالتان ممكنتان:

- يوجد شعاع تكامل مشترك وحيد

- يوجد عدة أشعة تكامل مشترك

إذا كنا في الحالة الأولى أي وجود شعاع تكامل مشترك واحد، نستخدم طريقة التقدير المقترحة من طرف(1987) Engle and Granger. كخطوة أولى، نقدر بطريقة المربعات الصغرى العادية OLS العلاقة طويلة المدى ثم حساب البواقي $\hat{\varepsilon}_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_{t1} - \hat{\beta}_2 X_{t2} - ... - \hat{\beta}_k X_{tk}$ وفي الخطوة الثانية، يتم تقدير العلاقة قصيرة المدى (النموذج الديناميكي –الحركي–) بطريقة OLS:

$$\nabla Y_t = \alpha_1 \nabla X_{t1} + \alpha_2 \nabla X_{t2} + ... + \alpha_k \nabla X_{tk} + \gamma_1 \hat{\varepsilon}_{t-1} + u_t$$

يجب أن يكون المعامل γ_1 معنويا سالبا. في أغلب الأحيان، لا يوجد شعاع تكامل مشترك واحد وعليه طريقةEngle and Granger غير صالحة. في هذه الحالة نحصل على مقدرات غير متسقة وغير كفؤة مهما تكن أشعة التكامل المشترك. إذن نحن مضطرون إلى استخدام الصيغة الشعاعية لنموذج تصحيح الخطأ Vector Error Correction Model.

2.3.3. نموذج تصحيح الخطأ المتعدد VECM:

لتكن السيرورة VAR(2) ذات k متغير على الشكل المصفوفي:

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \varepsilon_t$$

حيث:

Y_t : شعاع ذو بعد $(k \times 1)$ يتضمن المتغيرات $Y_{1,t},....,Y_{k,t}$

Φ_0 :شعاع ذو بعد $(k \times 1)$

Φ_i :مصفوفة المعاملات التي عناصرها $(k \times k)$.

يمكن كتابة النموذج باستعمال الفروقات من الدرجة الأولى:

$$Y_t - Y_{t-1} = \Phi_0 + \Phi_1 Y_{t-1} - Y_{t-1} + \Phi_2 Y_{t-2} + \varepsilon_t$$

أي:

$$\nabla Y_t = \Phi_0 + (\Phi_1 - I) Y_{t-1} + \Phi_2 Y_{t-2} + \varepsilon_t \quad ...... (1).$$

عند إضافة و طرح $\Phi_1 Y_{t-2} - Y_{t-2}$ لا يتغير شيء في المعادلة، أي:

$$\nabla Y_t = \Phi_0 + (\Phi_1 - I) Y_{t-1} + \Phi_1 Y_{t-2} - Y_{t-2} - \Phi_1 Y_{t-2} + Y_{t-2} + \Phi_2 Y_{t-2} + \varepsilon_t$$

$$\nabla Y_t = \Phi_0 + \Phi_1 Y_{t-1} - Y_{t-1} + \Phi_1 Y_{t-2} - Y_{t-2} - \Phi_1 Y_{t-2} + Y_{t-2} + \Phi_2 Y_{t-2} + \varepsilon_t \quad ....(2)$$

نضع: $\nabla Y_{t-1} = Y_{t-1} - Y_{t-2}$ ومن تم نقوم بتجميع العبارات في ∇Y_{t-1} وبعد التبسيط، نحصل على:

$$\nabla Y_t = \Phi_0 + (\Phi_1 - I)\nabla Y_{t-1} + (\Phi_1 + \Phi_2 - I)Y_{t-2} + \varepsilon_t \quad ....(3)$$

في هذه الأثناء، من أجل إعطاء علاقات التكامل المشترك الذي يربط بين المتغيرات الأصلية ذات فجوة زمنية واحدة نبحث على كتابة المعادلة (2) بدلالة Y_{t-1} أو في المعادلة (1) يمكن إضافة و طرح $\Phi_2 Y_{t-2}$ وبعد التبسيط، نحصل على:

$$\nabla Y_t = \Phi_0 - \Phi_2 \nabla Y_{t-1} + (\Phi_1 + \Phi_2 - I)Y_{t-1} + \varepsilon_t \quad ....(4)$$

أو:

$$\nabla Y_t = \Phi_0 + B_1 \nabla Y_{t-1} + \pi Y_{t-1} + \varepsilon_t \quad ....(5)$$

مع: $\Phi_2 = -B_1$ و $\pi = \Phi_1 + \Phi_2 - I$

يمكن تعميم هذه النتيجة على نموذج VAR(p) ذي k متغير:

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p} + \varepsilon_t$$

و كتابة النموذج على شكل فروقات من الدرجة الأولى بطريقتين:

$$\nabla Y_t = \Phi_0 + (\Phi_1 - I)\nabla Y_{t-1} + (\Phi_2 + \Phi_1 - I)\nabla Y_{t-2} + ... + (\Phi_{p-1} + ... + \Phi_2 + \Phi_1 - I)\nabla Y_{t-p+1} + \pi Y_{t-p} + \varepsilon_t$$

أو:

$$\nabla Y_t = \Phi_0 + B_1 \nabla Y_{t-1} + B_2 \nabla Y_{t-2} + ... + B_{p-1} \nabla Y_{t-p+1} + \pi Y_{t-1} + \varepsilon_t$$

المصفوفات B_i هي دوال للمصفوفات Φ_i و $\pi = \sum_{i=1}^{p} \Phi_i - I$.

يمكن كتابة المصفوفة π على الشكل $\pi = \alpha\beta'$ حيث أن الشعاع α يعبر عن قوة الجذب (الرجوع) نحو التوازن طويل المدى و β هو الشعاع الذي عناصره تعبر عن معاملات العلاقات طويلة المدى للمتغيرات. إذن كل توليفة خطية تعبر عن علاقة التكامل المشترك.

إذا كانت كل عناصر المصفوفة π معدومة (رتبة هذه المصفوفة تساوي الصفر و $I = \Phi_1 + \Phi_2 + ... + \Phi_{p-1}$)، فإنه لا يمكن تبني صياغة VECM. أما إذا كانت رتبة المصفوفة π تساوي k، فهذا يستلزم أن كل المتغيرات هي *I(0)* ومشكل التكامل المشترك لا يُطرح في هذه الحالة، أي أن تقدير نموذج VAR على السلاسل الأصلية هو نفسه التقدير على السلاسل ذات الفروقات.

إذا كانت رتبة المصفوفة π والتي نرمز لها بـ r. محصورة بين 1 و $k-1$ (أي $1 \le r \le k-1$)، فهذا يعني أنه يوجد r علاقة تكامل مشترك ونموذج VECM يعرف بالمعادلة التالية:

$$\nabla Y_t = \Phi_0 + B_1 \nabla Y_{t-1} + B_2 \nabla Y_{t-2} + ... + B_{p-1} \nabla Y_{t-p+1} + \alpha \hat{\varepsilon}_{t-1} + \varepsilon_t$$

مع: $\hat{\varepsilon}_t = \beta' Y_t$.

3.3.3. اختبار علاقة التكامل المشترك:

لتحديد عدد علاقات التكامل المشترك، اقترح Johansen (1988) اختبارا يعتمد على القيم الذاتية لمصفوفة يتم حسابها بإتباع الخطوتين التاليتين:

الخطوة الأولى: حساب البواقي $\hat{v}_t$; $\hat{u}_t$ انطلاقا من النموذجين التاليين:

$$\nabla Y_t = \hat{\Phi}_0 + \Phi_1 \nabla Y_{t-1} + \hat{\Phi}_2 \nabla Y_{t-2} + ... + \hat{\Phi}_p \nabla Y_{t-p} + u_t$$

$$Y_{t-1} = \hat{\Phi}_0 + \Phi_1 \nabla Y_{t-1} + \hat{\Phi}_2 \nabla Y_{t-2} + ... + \hat{\Phi}_p \nabla Y_{t-p} + v_t$$

مع: $$Y_t = \begin{pmatrix} Y_{1,t} \\ Y_{2,t} \\ . \\ . \\ . \\ Y_{k,t} \end{pmatrix}$$

u_t و v_t هي مصفوفات البواقي ذات بعد (k,T) حيث k هو عدد المتغيرات و T عدد المشاهدات.

<u>الخطوة الثانية</u>: حساب مصفوفات التباين- التباين المشترك التي تسمح بحساب القيم الذاتية

نقوم بحساب أربع مصفوفات ذات بعد (k,k) انطلاقا من بواقي التقدير $\hat{u}_t$ و $\hat{v}_t$:

$$\hat{\Sigma}_{uu} = \frac{1}{T}\sum_{t=1}^{T} u_t u_t^{'}$$

$$\hat{\Sigma}_{vv} = \frac{1}{T}\sum_{t=1}^{T} v_t v_t^{'}$$

$$\hat{\Sigma}_{uv} = \frac{1}{T}\sum_{t=1}^{T} u_t v_t^{'}$$

$$\hat{\Sigma}_{vu} = \frac{1}{T}\sum_{t=1}^{T} v_t u_t^{'}$$

ثم يتم الحصول على k قيمة ذاتية للمصفوفة M ذات بعد (k,k) المحسوبة كما يلي:

$$\hat{\Sigma}_{vv}^{-1}\hat{\Sigma}_{vu}\hat{\Sigma}_{uu}^{-1}\hat{\Sigma}_{uv}$$

انطلاقا من هذه القيم الذاتية، نقوم بحساب الإحصائية $\left(\lambda_{trace} = -T\sum_{i=r+1}^{k}\ln(1-\lambda_i)\right)$ مع λ_i القيمة الذاتية رقم i للمصفوفة M، k عدد المتغيرات و r رتبة المصفوفة.

تنبع هذه الإحصائية قانونا احتماليا يشبه إلى حد بعيد توزيع χ^2 مجدولا بالاستعانة بعملية محاكاة قام بها Johansen and Juselius (1990). يكون اختبار Johansen على الشكل التالي:

- رتبة المصفوفة π تساوي الصفر ($r=0$)، أي $H_0 : r = 0$ ضد الفرضية $H_1 : r > 0$. إذا رفضنا H_0، نمر إلى الاختبار الموالي (إذا كانت الإحصائية أكبر تماما من القيمة الحرجة لـ Johansen and Juselius ، فإننا نرفض H_0).
- رتبة المصفوفة π تساوي 1 ($r=1$)، أي $H_0 : r = 1$ ضد الفرضية $H_1 : r > 1$. إذا رفضنا H_0، نمر إلى الاختبار الموالي (إذا كانت الإحصائية أكبر تماما من القيمة الحرجة لـ Johansen and Juselius ، فإننا نرفض H_0).

- رتبة المصفوفة π تساوي 2 ($r=2$)، أي $H_0: r=2$ ضد الفرضية $H_1: r>2$. إذا رفضنا H_0، نمر إلى الاختبار الموالي (إذا كانت الإحصائية أكبر تماما من القيمة الحرجة لـ Johansen and Juselius ، فإننا نرفض H_0) وهكذا.

إذا رفضنا H_0 في نهاية المطاف واختبرنا بعدها الفرضية $H_0: r=k-1$ ضد $H_1: r=k$ وقمن برفض H_0، فإن رتبة المصفوفة هي $r=k$ وفي هذه الحالة لا يوجد علاقة تكامل مشترك باعتبار أن المتغيرات هي *I(0)*.

للقيام بهذا الاختبار، اقترح Johansen (1988) خمس صيغ تتعلق بأشعة التكامل:

1. في حالة عدم وجود الاتجاه العام الخطي في المعطيات:

أ. غياب الاتجاه الخطي في السلاسل وغياب الحد الثابت في علاقات التكامل المشترك.

ب. غياب الاتجاه الخطي في السلاسل ووجود الحد الثابت في علاقات التكامل المشترك.

2. في حالة وجود الاتجاه العام الخطي في المعطيات:

ج. وجود الاتجاه الخطي في السلاسل و الحد الثابت في علاقات التكامل المشترك.

د. وجود الاتجاه الخطي في السلاسل و في علاقات التكامل المشترك.

3. في حالة وجود اتجاه عام كثير حدود من الدرجة الثانية في المعطيات:

هـ .. وجود اتجاه عام كثير حدود من الدرجة الثانية في السلاسل واتجاه خطي في علاقات التكامل المشترك.

إن اختيار إحدى الصيغ الخمسة يتوقف بالدرجة الأولى على الشكل الرياضي للاتجاه العام، فمثلا يمكن الاستعانة بالشكل البياني للسلاسل الذي يسمح بتحديد الشكل المناسب.

يسمح اختبار Johansen بتحديد عدد علاقات التكامل المشترك ولكن لا يبين ما هي المتغيرات المتكاملة Cointegrated variables. غير أن هناك اختبار قيود يسمح بتحديد المتغيرات المتكاملة[1].

مثال 6[2]:

لتكن المعطيات الخاصة بالمتغيرات Y_{1t}, Y_{2t}, Y_{3t}. المطلوب اختبار التكامل المشترك بين هذه المتغيرات وتقدير نموذج VAR أو نموذج VECM. ليكن الجدول التالي:

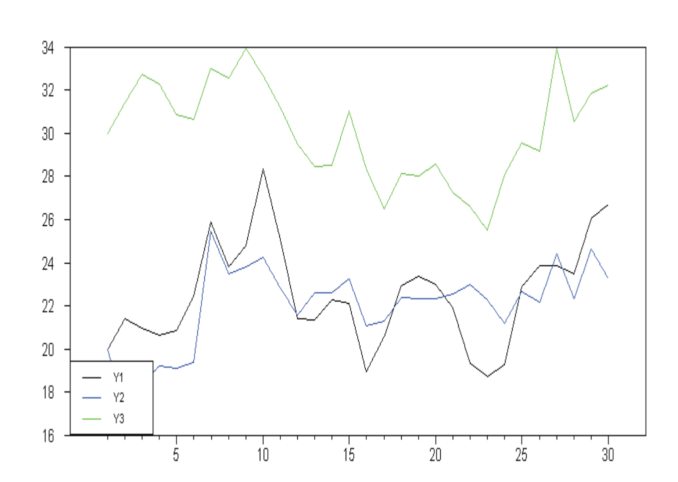
الجدول (5): المعطيات المشاهدة Y_{1t}, Y_{2t}, Y_{3t}

Y_{3t}	Y_{2t}	Y_{1t}	المشاهدات
30.00	20.00	20.00	1
31.40	17.05	21.39	2
32.71	18.32	20.95	3
32.32	19.22	20.62	4
30.87	19.12	20.88	5
30.67	19.36	22.42	6
33.03	25.44	25.89	7
32.59	23.46	23.83	8
33.91	23.81	24.78	9
32.65	24.27	28.37	10
31.20	22.85	25.06	11
29.48	21.59	21.40	12
28.49	22.62	21.33	13
28.52	22.62	22.29	14
31.02	23.25	22.14	15
28.38	21.08	18.97	16
26.52	21.32	20.59	17
28.12	22.38	22.96	18
28.04	22.36	23.36	19
28.57	22.32	22.98	20
27.28	22.58	21.91	21
26.61	22.98	19.33	22
25.49	22.27	18.70	23
28.10	21.21	19.30	24
29.57	22.69	22.86	25
29.19	22.15	23.85	26
33.91	24.40	23.85	27
30.53	22.36	23.46	28
31.84	24.61	26.04	29
32.22	23.25	26.70	30

1- Hamilton (1994), p. 648-651 23.83
2- Bourbonnais (2003), p. 295

نمثل بيانيا هذه السلاسل:

الشكل رقم (5): التمثيل البياني للسلاسل



لتقدير النموذج VAR، نطبق طريقة المربعات الصغرى العادية على كل معادلة. نختار النموذج الأمثل بالاستعانة بالمعيارين Akaike و Schwarz من أجل رتبة (درجة) تتغير من 0 إلى 3، فاقتراح درجات أكبر يؤدي إلى فقدان عدة مشاهدات وكما نعلم، عدد المشاهدات في هذه السلاسل صغير. قدرنا ثلاث نماذج مختلفة واخترنا نموذج VAR(1) لأنه يصغر المعيارين AIC و Schwarz، استعملنا لهذا الغرض برمجية RATS 5.04.

نقوم إذن بتطبيق اختبار Johansen على نموذج VAR(1) تحت فرضيتين:

الفرضية الأولى: وجود حد ثابت في العلاقة طويلة المدى وليس في المعطيات (بدون اتجاه عام تحديدي)

نختبر الفرضية التالية: رتبة المصفوفة π تساوي الصفر ($r = 0$)، أي $H_0 : r = 0$ ضد الفرضية $H_1 : r > 0$.

القيم الذاتية الثلاثة للمصفوفة π المقدرة بطريقة المعقولية العظمى هي: $\lambda_1 = 0.605$ ؛ $\lambda_2 = 0.257$ ؛ $\lambda_3 = 0.139$. في هذه الحالة، نحسب إحصائية Johansen $\lambda_{trace} = -T\sum_{i=r+1}^{k}\ln(1-\lambda_i)$ من أجل $r = 0$. لدينا:

$$\lambda_{trace} = -T \times [\ln(1-\lambda_1) + \ln(1-\lambda_2) + \ln(1-\lambda_3)]$$
$$\lambda_{trace} = -28 \times [0.928 + 0.297 + 0.150] = 38.50$$

القيم الحرجة تساوي 34.31 عند مستوى معنوية 5% و 41.07 عند 1%. نرفض إذن الفرضية H_0 ؛ رتبة المصفوفة لا تساوي الصفر وهذا يعني أن السلاسل غير مستقرة.

نختبر الفرضية التالية: رتبة المصفوفة π تساوي 1 $(r = 1)$، أي $H_0 : r = 1$ ضد الفرضية $H_1 : r > 1$. لدينا:

$$\lambda_{trace} = -T \times [\ln(1-\lambda_2) + \ln(1-\lambda_3)]$$
$$\lambda_{trace} = -28 \times [0.297 + 0.150] = 12.51$$

القيم الحرجة تساوي 19.96 عند مستوى معنوية 5% و 24.60 عند 1%. إذن لا يمكن رفض الفرضية H_0 ؛ نعتبر أن رتبة المصفوفة تساوي 1 وهذا يعني قبول فرضية علاقة التكامل المشترك.

الفرضية الثانية: وجود حد ثابت في العلاقة طويلة المدى وفي المعطيات

تظهر النتائج كما يلي:

الجدول (6): نتائج الاختبار

القيمة الحرجة عند %1	القيمة الحرجة عند %5	λ_{trace}	القيمة الذاتية
35.65	29.68	37.388	0.603
20.04	15.41	11.469	0.229
6.65	3.76	4.175	0.138

من خلال النتائج المبينة في الجدول، نلاحظ أن رتبة المصفوفة π لا تساوي الصفر (السطر الأول) ولكن من جهة أخرى لا يمكن رفض الفرضية H_0 لا عند مستوى معنوية 5% ولا عند 1% (السطر الثاني)، في هذه الحالة رتبة المصفوفة تساوي 1. هناك علاقة وحيدة محققة للتكامل المشترك.

ووفقا للنتائج المتحصل عليها، نقوم بتقدير نموذج تصحيح الخطأ المتعدد VECM في كلتا الحالتين (بوجود ثابتة في المعطيات أو بدونها) وفي ظل وجود علاقة وحيدة للتكامل المشترك.

لقد تم رفض الصيغة الأولى (الفرضية الأولى) لنموذج VECM باعتبار أن الحدود الثابتة للمعادلات الثلاثة ليست لها معنوية إحصائية وعليه التقدير النهائي لهذا النموذج على 28 مشاهدة يعرف كما يلي:

$$\underset{}{\nabla Y_{1t}} = \underset{(2.64)}{-0.85} \times \left(Y_{1t-1} - \underset{(8.9)}{0.87} Y_{2t-1} - \underset{(7.1)}{0.54} Y_{3t-1} + \underset{(4.19)}{12.92}\right)$$

$$+ \underset{(2.82)}{0.77} \nabla Y_{1t-1} - \underset{(2.61)}{0.71} \nabla Y_{2t-1} - \underset{(1.07)}{0.26} \nabla Y_{3t-1}$$

$$\nabla Y_{2t} = \underset{(1.18)}{0.32} \times \left(Y_{1t-1} - \underset{(8.9)}{0.87} Y_{2t-1} - \underset{(7.1)}{0.54} Y_{3t-1} + \underset{(4.19)}{12.92}\right)$$

$$+ \underset{(0.43)}{0.10} \nabla Y_{1t-1} - \underset{(1.44)}{0.34} \nabla Y_{2t-1} - \underset{(0.39)}{0.08} \nabla Y_{3t-1}$$

$$\nabla Y_{2t} = \underset{(0.03)}{0.0089} \times \left(Y_{1t-1} - \underset{(8.9)}{0.87} Y_{2t-1} - \underset{(7.1)}{0.54} Y_{3t-1} + \underset{(4.19)}{12.92}\right)$$

$$+ \underset{(2.16)}{0.53} \nabla Y_{1t-1} - \underset{(1.86)}{0.46} \nabla Y_{2t-1} - \underset{(1.13)}{0.25} \nabla Y_{3t-1}$$

القيم التي بين قوسين هي قيم ستيودنت.

نلاحظ في المعادلة الأولى أن المعامل الذي يعبر عن قوة الجذب (الرجوع) نحو التوازن طويل المدى معنويا سالب ولمعاملات العلاقة طويلة المدى معنوية إحصائية، فهي تختلف كلها معنويا عن الصفر، أما في المعادلتين الأخيرتين هناك بعض المعاملات التي ليست لها معنوية إحصائية، فهذه الأخيرة لا تؤثر بشكل أو آخر على المعنوية الكلية لنموذج VECM.

تخضع بواقي التقدير لكل معادلة لسيرورة التشويش الأبيض أي تتميز بالاستقرار طبقا لإحصائية Ljung-Box:

الجدول (7): نتائج اختبار Ljung-Box

	الإحصائية $Q(12)$	الاحتمالات الحرجة
المعادلة الأولى	5.87	0.92
المعادلة الثانية	4.60	0.97
المعادلة الثالثة	5.99	0.92

تبقى إحصائيات Ljung-Box أقل تماما من القيمة الحرجة لتوزيع χ^2 باعتبار أن الاحتمالات الحرجة أكبر تماما من 0.05 وعليه نقبل الفرضية H_0. يمكن القول في الأخير أن صيغة VECM مقبولة وصحيحة.

الفصل الثامن

نماذج ARCH وتطبيقاتها المالية

الفصل الثامن

نماذج ARCH وتطبيقاتها المالية

تطرقنا في الفصل السادس إلى مختلف النماذج الخطية للسلاسل الزمنية التي ساهمت بدور كبير في نمذجة الكثير من الظواهر الاقتصادية، واستطاعت أن تُعطي لعدة نظريات صورة رياضية تُساعد على التنبؤ بالقيم المستقبلية. إلا أن ما يؤخذ على هذه الصيغ الخطية أنها لا تستطيع أن تُتَرجم الصفة الحركية لهذه الظواهر، وهذا ما أدى إلى عرقلة تطور عدة جوانب النمذجة في السلاسل الزمنية، ففرضية الخطية التي تتصف بها هذه النماذج تستلزم أن تتميز المكونات الزمنية بوقت واحد، إضافة إلى ذلك أن ثبات السيرورة ARMA، لا يسمح بأخذ الميكانيزمات غير المتناظرة بعين الاعتبار، أما فيما يخص نموذج الانحدار الذاتي AR(p)، فهو يُفسِّر القيمة الحالية للسلسلة بدلالة القيم الماضية، ومن ثم لا يستغل استغلالا كاملا للمعلومات الموجودة في السلسلة. في سنة 1982، اقترح Engle (1982) نموذجا غير خطي، يعبر عن الانحدار الذاتي الذي يتضمن تباينا شرطيا غير متجانس باستعمال معلومات سابقة، يسمى بنموذج ARCH.

وفي هذا الفصل سنلقي الضوء على أهم نماذج ARCH و سنرى مدى نجاعتها في التنبؤ بالقيم المستقبلية، والصيغ الحديثة الناتجة منها.

1. مفاهيم أساسية

إن دور صفة "عدم التأكد" في تحديد حركية سلوك مختلف المتغيرات الاقتصادية الحديثة، خاصة في المسائل المالية، جعل النظريات الاقتصادية القياسية تعطيه قدرا من الأهمية، بدءاً باستخدام المتوسط الشرطي بدلا من المتوسط غير الشرطي في نماذج ARMA، هذه الصفة الإضافية من شأنها أن تساهم في تحسين التنبؤات الناتجة عن هذه النماذج المختلطة، وللتفرقة بين هذين المفهومين، نعتبر السيرورة التالية:

$$AR(1):\quad Y_t = \phi_1 Y_{t-1} + \varepsilon_t$$

حيث ε_t هو تشويش أبيض و المتوسط (التوقع) الشرطي يعطى بالعلاقة $\phi_1 Y_{t-1}$ $= E(Y_t \mid Y_{t-1}, Y_{t-2}, ...)$ بينما يكون المتوسط غير الشرطي معدوما.

بعد ذلك تطورت هذه الفكرة لتشمل العزوم من الدرجة الثانية، حيث أشار Engle سنة 1982 إلى أهمية استعمال مفهوم التباين الشرطي بدلا من التباين غير الشرطي في تحسين القيم التنبؤية، لأنه بينما يبقى هذا الأخير ثابتا بتغير الزمن، فإن التباين الشرطي يمكن أن يُترجِم العلاقة بين المشاهدة Y_t، والمشاهدات السابقة Y_{t-i}. فإذا أخذنا المثال السابق، يكون التباين الشرطي للسيرورة AR(1) من الشكل:

$$\mathrm{var}(Y_t \mid Y_{t-1}, Y_{t-2}, ...) = E\left(\left[Y_t - E\left(Y_t / Y_{t-1}, Y_{t-2}, ...\right)\right]^2 \mid Y_{t-1}, Y_{t-2}, ...\right)$$

بينما يكون التباين غير الشرطي كما يلي:

$$\mathrm{var}(Y_t) = \sigma_\varepsilon^2 / \left(1 - \phi_1^2\right)$$

كل هذه المبادئ كانت بساطا يُفرَش لصياغة النماذج ARCH، وهي نماذج انحدار ذاتي ذات تباين شرطي غير متجانس. حيث أراد Engle (1982) من خلالها سد النقص الذي كانت تعاني منه نماذج ARMA السابقة، خاصة في السلاسل المالية التي تتميز بـ سرعة التقلبات The volatility المرتبطة بالزمن.

1.1. مشكل عدم تجانس تباينات الأخطاء Heteroscedasticity Problem:

إن معظم النماذج الكلاسيكية التي تطرقنا إليها في الفصول السابقة، ترتكز على فكرة أساسية تتمثل في أن متوسط الأخطاء معدوم و تباينها ثابت مع تغير الزمن وأنها مستقلة عن بعضها البعض أي:

$$E(\varepsilon_t) = 0 \quad , \forall t = 1,....,T$$

$$Var(\varepsilon_t) = E(\varepsilon_t^2) = \sigma^2 \quad , \forall t = 1,.....,T$$

$$Cov(\varepsilon_t, \varepsilon_{t'}) = E(\varepsilon_t \varepsilon_{t'}) = 0 \quad , \forall t \neq t' \qquad t, t' = 1,....,T$$

وبإسقاط هذه الفرضيات فإن تقدير مصفوفة التباين والتباين المشترك يصبح صعبا، لأن الأخطاء ستكون غير متجانسة ومترابطة فيما بينها، مما يقلل من نجاعة النماذج المقدرة. و

في هذا الإطار كانت هناك العديد من الأعمال المقدمة والحلول المقترحة حول مصفوفة التباين المستحدثة، أدت بدورها إلى جملة من التساؤلات، من بينها[1]:

- ❖ كيف نبني نموذجا رياضيا يسمح بدراسة الشكل المقترح؟.
- ❖ كيف نقوم بتقدير معالم هذا النموذج؟.
- ❖ كيف نكتشف وجود شكل معين؟.

إن التفكير البسيط يميل إلى تكبير حجم العينة T عند تقدير مصفوفة التباين وهذا من أجل الحصول على تقديرات متقاربة، غير أن هذه الطريقة لا تحل المشكلة إلا بصفة جزئية فقط، كونها تؤدي إلى تكبير عدد المعالم المقدرة، من أجل ذلك اقترح الباحثون جملة من الأفكار، ليكن على سبيل المثال نموذج الانحدار الذاتي AR(1) المعرف بالشكل [2]:

$$\varepsilon_t = \rho\varepsilon_{t-1} + \mu_t$$

حيث μ_t يمثل تشويشا أبيض:

$$E(\mu\mu') = \begin{bmatrix} \sigma_\mu^2 & 0 & \cdots & 0 \\ 0 & \sigma_\mu^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_\mu^2 \end{bmatrix}$$

في هذا النوع من النماذج تأخذ مصفوفة التباين–التباين المشترك لـ ε شكلا خاصا:

$$E(\varepsilon\varepsilon') = \begin{bmatrix} E(\varepsilon_1^2) & E(\varepsilon_1\varepsilon_2) & \cdots & E(\varepsilon_1\varepsilon_{t'}) & \cdots & E(\varepsilon_1\varepsilon_T) \\ E(\varepsilon_2\varepsilon_1) & E(\varepsilon_2^2) & \cdots & E(\varepsilon_2\varepsilon_{t'}) & \cdots & E(\varepsilon_2\varepsilon_T) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ E(\varepsilon_i\varepsilon_1) & E(\varepsilon_i\varepsilon_2) & \vdots & E(\varepsilon_i\varepsilon_{t'}) & \vdots & E(\varepsilon_i\varepsilon_T) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ E(\varepsilon_T\varepsilon_1) & E(\varepsilon_T\varepsilon_2) & \cdots & E(\varepsilon_T\varepsilon_{t'}) & \cdots & E(\varepsilon_T^2) \end{bmatrix}$$

1- Terreza and Zatout, p. 39.
2- Johnston (1988), Tome 2, p. 362.

حيث أن:

❖ في حالة $t = t'$: العناصر القطرية كلها متساوية:

$$E\left(\varepsilon_1^2\right) = ... = E\left(\varepsilon_T^2\right) = \sigma_\varepsilon^2 = \frac{\sigma_\mu^2}{1-\rho^2}$$

❖ في حالة $t \neq t'$: العناصر غير القطرية كلها متساوية لعدد أسي متقـ ـارب:

$$E\left(\varepsilon_t \varepsilon_{t-k}\right) = \rho^k \sigma_\varepsilon^2$$

نلاحظ أن في كلتا الحالتين من أجل أخطاء مرتبطة من الشكل AR(1)، يكفي أن نقدر المعلمتين σ^2 و ρ، فمن بين أهم أسباب وجود عدم تجانس التباين في السلسلة، هو الحالة التي تكون المشاهدات في شكل مجموعات غير متجانسة، فعلى سبيل المثال نجد الإنفاق الأسري يوجه عادة إلى السلع الضرورية عند الأسر ضعيفة الدخل، في حين أن الأسر الغنية سيكون توزيع نفقاتها متذبذبا بين السلع الكمالية ذات السعر المرتفع والسلع الضرورية مما يؤدي إلى مشكلة عدم تجانس التباين في المعطيات المجمعة (من حيث أن تباينات الأخطاء التابعة للمجموعة الأولى أكبر نسبيا مما هي عليه في المجموعة الثانية)[1].

لحل مشكلة عدم تجانس تباين الأخطاء أُقترِحَت عدة أفكار وحلول، ترتكز في معظمها على إيجاد تباين يتطور مع الزمن، ومن بينها إدخال متغيرات جديدة X_t تُفسر هذا التطور، إضافة إلى ذلك كون التباين ثابتا في كل مجموعة أو فئة و يؤخذ التباين أو الانحراف المعياري كأنه دالة خطية لمتغيرات خارجية ويفترض هنا أن المتغير الداخلي يكون مستقلا عن تغير التباين.

ومن خلال دراسة لمعدلات التضخم في المملكة المتحدة سنة 1982. اقتـ ـرح Engle (1982) استخدام المتغيرة العشوائية لتفسير عدم التجانس الشرطي، وهذا ما تمخض عنه ما يسمى بنماذج الانحدار الذاتي ذات التباين الشرطي غير المتجانس ARCH.

1- Pindyck and Rubenfled (1981), p. 139.

2.1. أثر استخدام التوزيع الشرطي على التوقع:

لتحليل هذا الأثر ندرس كمثال حالة سيرورة ماركوفية من الدرجة الأولى[1]:

$$AR(1) \quad : \quad \varepsilon_t = \phi_1 \varepsilon_{t-1} + \mu_t$$

حيث $\mu_t \sim N(0,\sigma^2)$ و $P(\varepsilon_t \mid s<t) = P(\varepsilon_t \mid \varepsilon_{t-1})$

بافتراض أن السيرورة مستقرة، أي: $|\phi_1|<1$، إذن:

$$\varepsilon_t \mid \varepsilon_{t-1} \sim N(\phi_1\varepsilon_{t-1}, \sigma^2) \quad \text{و} \quad \varepsilon_t \sim N\left(0, \frac{\sigma^2}{1-\phi_1^2}\right)$$

معنى هذا أن استخدام التوزيع الشرطي يمكن أن يحسن نوعية مجال التوقع، حيث يُوظف المتوسط $\phi_1\varepsilon_{t-1}$، هذا من جهة، ومن جهة أخرى أن الانحراف قد قل من: $\pm\frac{\sigma}{\sqrt{1-\phi_1^2}}$ إلى $\pm\sigma$.

لكن الشيء الملاحظ أن ذاكرة السيرورة لا تظهر في انحراف التوقع سواءً أكان شرطيا أم لا، أي أن قيم التباين غير مرتبطة بالقيم السابقة للسيرورة وعليه لا يوجد أي تحسن في حدود مجال التوقع. من هنا تظهر أهمية التعديلات التي قام بها Engle (1982)، حيث قدم نموذج التباين العشوائي بطريقة داخلية. كما قام بإدراج المشاهدات السابقة للسيرورة في شكل انحدار ذاتي لمربعات الأخطاء.

2. التحاليل النظرية حول نماذج ARCH/GARCH

1.2. صياغة نموذج ARCH(p) و خصائصه:

تعرف السيرورة ARCH "*Autoregressive Conditional Heteroscedastic*" كتشويش أبيض يخضع للتوزيع الطبيعي η_t مضروبة من أجل كل فترة بمتغير عشوائي $h_t^{1/2}$ الذي يرتبط خطيا بالقيم الماضية للسيرورة:

1- Hamidi, Khenouse and Zatout (1998).

$$\varepsilon_t = \eta_t \times h_t^{1/2},$$
$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2$$
$$\eta_t \to N(0,1)$$

يمكن التعبير عن ε_t بدلالة I_t، كمية المعلومات المتاحة في الفترة t والتوزيع الشرطي لـ ε_t طبيعي مركز ذي تباين h_t:

$$E(\varepsilon_t | I_{t-1}) = 0$$
$$\mathrm{var}(\varepsilon_t | I_{t-1}) = h_t$$

إضافة إلى ذلك، يمكن صياغة ε_t^2 على شكل سيرورة AR(p). نضع:

$$v_t = \varepsilon_t^2 - h_t$$

مع:
$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2$$

أي:
$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + v_t$$

حيث لـ v_t متوسط و تباين مشترك معدوم ولكن تباين غير ثابت.

يمكن الحصول على نموذج الانحدار ARCH و ذلك بافتراض أن متوسط ε_t توليفة خطية للمتغيرات الخارجية و الداخلية المدرجة في شعاع المعلومات I_{t-1} مضروبا بشعاع معالم مجهولة:

$$\varepsilon_t | I_{t-1} \to N(x_t \beta, h_t)$$
$$h_t = h(\eta_{t-1}, \eta_{t-2}, ..., \eta_{t-p}, \alpha)$$
$$\eta_t = \varepsilon_t - x_t \beta$$

تملك هذه العبارة خصائص مهمة في التطبيقات القياسية و ذلك باعتبار أن "عدم التأكد" المتعلق بالتنبؤ يتغير بتغير الفترات و ليس فقط مع أفق التنبؤ و الأخطاء العشوائية تتجمع عادة على شكل أخطاء مرتفعة متبوعة بأخطاء ضعيفة. إن الصيغة الرياضية لـ ARCH، حيث التباين يرتبط بالزمن و الأخطاء السابقة، تسمح بالأخذ بعين الاعتبار هذه الظاهرة. إذا كانت المعاملات α_i كلها موجبة و كبيرة نسبيا، يوجد ما يسمى بالصمود

" Persistence" على مستوى التقلبات "Volatility"، نشاهد إذن فترات تطاير قوية تتبعها فترات تطاير ضعيفة.

إضافة إلى ذلك، إذا كنا في النظرية المالية أو النقدية، نشير هنا إلى أهمية نمذجة ARCH. محافظ السندات مثلا هي دوال لمتوسط و تباين المردودية. كل تعديل للطلب على السند يجب أن يكون مرتبطا بتغيرات المتوسطات و التباينات المتوقعة للمردودية، ففي هذه الحالة عندما يتبع المتوسط نموذج انحدار عادي، يكون التباين ثابتا و هذا ما يتناقض مع هذه الحالة.

يسمح هذا النوع من النماذج بنمذجة حركية (أو ديناميكية) للتطاير و يوفق بين الحركية الاحتمالية و التمثيل الهيكلي للظاهرة المدروسة و يساعدنا على تحليل تطاير الأصول المالية.

لكي يكون التباين الشرطي $\text{var}(\varepsilon_t \mid I_{t-1})$ موجبا و محدودا (أقل من ∞)، فينبغي أن تكون الشروط التالية على المعالم محققة:

$$\alpha_0 > 0, \alpha_1 \geq 0, \ldots\ldots, \alpha_p \geq 0$$

$$\sum_{i=1}^{p} \alpha_i < 1$$

يعرف مؤشر Kurtosis على أنه نسبة العزم المركز من الدرجة 4 على مربع العزم المركز من الدرجة 2، في حالة ARCH(1)، لدينا:

$$K = \frac{E(\varepsilon_t^4)}{[E(\varepsilon_t^2)]^2} = \frac{3(1-\alpha_1^2)}{1-3\alpha_1^2}$$

المقدار K أكبر تماما من 3 و هذا ما نلاحظه خاصة في السلاسل الزمنية المالية حيث أنها تحتوي على شكل توزيع مفلطح أي أكثر سمكا من التوزيع الطبيعي و هذا هو حال السيرورة ARCH التي لها توزيع مفلطح "Leptokurtic distribution".

مثال 1:

نقوم بعملية محاكاة للسيرورة ARCH التالية (اصطناع سلسلة تتبع سيرورة ARCH)

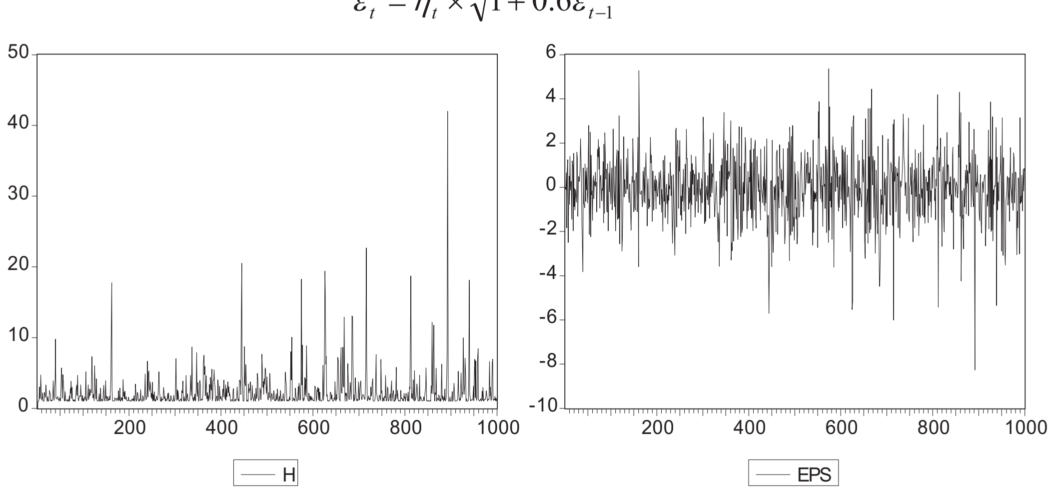
$$\varepsilon_t = \eta_t \times \sqrt{1+0.6\varepsilon_{t-1}^2}$$

نستخدم برمجية Eviews 5.0 لمحاكاة السيرورة مع $T = 1000$:

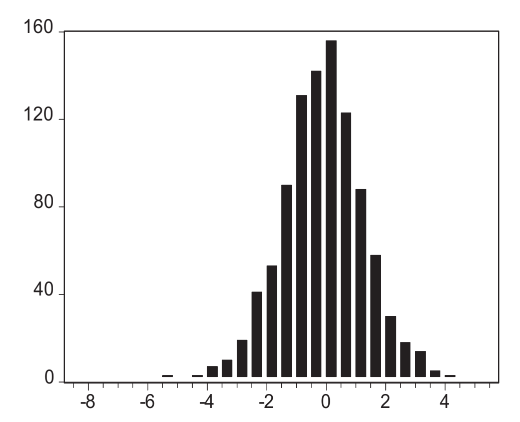
```
CREATE U 1000
GENR U = NRND
GENR H = 0
GENR H = 1+0.6*U(-1)*U(-1)*H(-1)
GENR EPS = SQR(H)*U
```

الشكل رقم (1): محاكاة الخطأ **الشكل رقم (2): محاكاة التباين الشرطي**

$$\varepsilon_t = \eta_t \times \sqrt{1+0.6\varepsilon_{t-1}^2}$$



الشكل رقم (3): توزيع بواقي التقدير



Series: EPS	
Sample 1 1000	
Observations 1000	
Mean	-0.045810
Median	-0.009909
Maximum	5.362981
Minimum	-8.265671
Std. Dev.	1.493371
Skewness	-0.313221
Kurtosis	4.839246
Jarque-Bera	157.3022
Probability	0.000000

نلاحظ من خلال الشكلين (1) و (2) أن الأخطاء العشوائية و التباين ليست ثابتة بل تتغير بتغير الزمن، فتغيرات الخطأ تعرف مراحل استقرار متبوعة بفترات تباين أكثر ارتفاعا كما نلاحظ أيضا من خلال الشكل (3) أن معامل Kurtosis أكبر تماما من 3 و هذا يعني أن للسيرورة ARCH توزيع مفلطح "Leptokurtic distribution".

للأخذ بعين الاعتبار حركية التباين الشرطي للأخط اء، عم م Bollerslev (1986) نمذجة سرعة التقلبات الشرطية Conditional Volatility ويسمى هذا النوع من النماذج بنم . . .وذج GARCH(p,q) " Generalized Autoregressive Conditional Heteroscedastic" و الذي يكتب رياضيا كما يلي:

$$Y_t = X\beta + \varepsilon$$

$$\varepsilon_t = \eta_t \times h_t^{1/2} \quad , \quad \eta_t \sim N(0,1)$$

$$h_t = \text{var}(\varepsilon_t \mid I_{t-1}) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}^2$$

$$\alpha_0 > 0 \, , \alpha_i \geq 0 \, , \beta_j \geq 0 \, , i = 1,...., p \, , \, j = 1,...., q$$

إن السيرورة GARCH(p,q) هي سيرورة ARCH من الدرجة (الرتبة) ∞، حيث أن المعالم تتناقص بوتيرة هندسية. تعتبر هذه السيرورة حلا بديلا، تحتفظ ببنية تباطؤ أكثر

بساطة و يعطي ذاكرة أكبر. يمكن أيضا صياغة هذه السيرورة على شكل نموذج ARMA الكلاسيكي، فهي كتابة أكثر استعمالا لمعالجة مشكل الاستقرارية. ليكن $v_t = \varepsilon_t^2 - h_t$، المعادلة تصبح:

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j} = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)h_t$$

ومنه:

$$[1 - \alpha(L) - \beta(L)]\varepsilon_t^2 = \alpha_0 + [1 - \beta(L)]v_t$$

كنتيجة لذلك، يمكن كتابة نموذج GARCH بصيغة $ARMA(\max(p,q),q)$ على مربع الأخطاء ε_t. تكون هذه السيرورة مستقرة "بصفة ضعيفة" إذا كان:

$$\alpha(1) + \beta(1) = \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$$

2.2. اختبار نموذج ARCH/GARCH:

ليكن النموذج الكلاسيكي $Y = X\beta + \varepsilon$ و لنفترض أن البواقي تخضع للسيرورة ARCH حيث $\varepsilon_t = \eta_t \times h_t^{1/2}$ مع $h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2$. يكون التباين الشرطي للأخطاء متجانسا إذا تحققت الفرضية $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_p = 0$. عكس ذلك، إذا كانت المعاملات α_i لا تنعدم كلها، فهذا يعني أن التباين الشرطي للأخطاء غير ثابت والأخطاء العشوائية تتبع نموذج ARCH(p) التي ينبغي تحديد درجته p.

يرتكز الاختبار إما على اختبار فيشر الكلاسيكي أو اختبار مضاعف لاغرانج LM. لتطبيق هذا الاختبار، لا بد أولا من حساب بواقي تقدير النموذج العام الكلاسيكي أو نموذج ARMA، $\hat{\varepsilon}_t$ ثم حساب مربعات البواقي $\hat{\varepsilon}_t^2$ وبعد ذلك القيام بتقدير انحدار $\hat{\varepsilon}_t^2$ على مربعات البواقي في الفترات السابقة، أي $\hat{\varepsilon}_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \hat{\varepsilon}_{t-i}^2$. نحسب معامل التحديد R^2 الخاص بالمعادلة الأخيرة ليتم بعد ذلك حساب إحصائية مضاعف لاغرانج $LM = T \times R^2$ (مع T عدد المشاهدات) الذي يخضع لتوزيع χ^2 بدرجة حرية p و نسبة

معنوية α. إذا كان $LM > \chi^2_\alpha(p)$، نرفض الفرضية H_0 أي نعتبر أنه يمكن قبول ARCH(p).

هناك طريقة أخرى تتمثل في التمثيل البياني لدالة الارتباط الذاتي لمربعات البواقي. إذا كانت بعض (أو معظم) معاملات الارتباط الذاتي تختلف معنويا عن الصفر بنسبة معنوية α (أي إذا كانت سلسلة مربعات البواقي غير مستقرة)، فإنه يمكن قبول فرضية عدم تجانس التباين الشرطي للأخطاء، أي أن الظاهرة المالية تخضع للسيرورة ARCH. نستعمل في هذا الصدد إحصائية Ljung-Box التي تم التطرق إليها في الفصل السادس.

يمكن أيضا اختبار نموذج ARCH ضد نمذجة GARCH، أي نختبر فرضية العدم $H_0 : \beta_1 = \beta_2 = ... = \beta_q = 0$ ضد الفرضية البديلة $H_1 : \exists \beta_j \neq 0$. من أجل ذلك، نقوم بحساب معامل التحديد R^2 الخاص بالمعادلة $\hat{h}_t = \hat{\alpha}_0 + \sum_{i=1}^{p} \hat{\alpha}_i \varepsilon^2_{t-i} + \sum_{j=1}^{q} \hat{\beta}_j h_{t-j}$ ونقارن إحصائية مضاعف لاغرانج بالقيمة الحرجة لتوزيع χ^2 بدرجة حرية q. إذا كانت هذه الإحصائية أكبر تماما من القيمة المجدولة لهذا التوزيع، فإن الأخطاء تخضع لنموذج GARCH(p,q) وينبغي في هذه الحالة تحديد p و q بتصغير معيار AIC أو Schwarz.

هناك اختبار آخر مكافئ لاختبار LM يرتكز على مقارنة نسبة المعقولية للنموذج المقيد ARCH(p) و غير المقيد GARCH(p,q). إذا كانت القيود موجودة $\beta_1 = \beta_2 = ... = \beta_q = 0$ فهذا يعني أن $L_c < L_{nc}$ حيث L_{nc} هي دالة المعقولية للنموذج غير المقيد و L_c للنموذج المقيد، أي أن $L_c / L_{nc} < 1$ أو بشكله اللوغاريتمي $\ln L_c - \ln L_{nc} < 0$. الفرق بين لوغاريتمات الدالة ينبغي أن يكون معنويا سالبا. نقوم بحساب الإحصائية $LR = -2(\ln L_c - \ln L_{nc})$ التي تخضع لتوزيع χ^2 بدرجة حرية q و التي تعبر عن عدد القيود. إضافة إلى ذلك، إذا كان LR أكبر من القيمة المجدولة لتوزيع χ^2 بنسبة معنوية α و درجة حرية q، نرفض الفرضية H_0 أي أن القيود ليست محققة و بعبارة أخرى نقبل النموذج GARCH(p,q).

3. التقدير والتنبؤ

هناك ثلاث طرق رئيسية لتقدير النماذج ذات أخطاء تتميز بخاصية عدم تجانس التباين الشرطي، ينتج عنها ثلاث أنواع من المقدرات هي:

- ❖ مقدرات من فئة المعقولية العظمى (MLE) Maximum Likelihood Estimators
- ❖ مقـدرات المعقوليـة العظمـى الزائفـة Pseudo-Maximum Likelihood Estimators (Pseudo-MLE)
- ❖ مقدرات عن طريق مرحلتين (GLS) Generalized Least Squares Method

يمكن تمثيل معظم نماذج ARCH بالشكل التالي.في هذا الإطار نأخذ النموذج المقدم من طرف (1992) Gouriéroux :

$$E(Y_t \mid Y_{t-1}, X_t) = m_t(Y_{t-1}, X_t, \theta) = m_t(\theta)$$
$$\text{var}(Y_t \mid Y_{t-1}, X_t) = h_t(Y_{t-1}, X_t, \theta) = h_t(\theta)$$

حيث θ هي مجموعة المعالم الداخلة في صيغة كل من المتوسط الشرطي والتباين الشرطي.

سنحاول تقديم بشكل موازي طريقة تقدير المعقولية العظمى ML تحت فرضية التوزيع الشرطي الطبيعي للبواقي، مع طريقة Pseudo-ML ، حيث نجد في الحالتين أن دالة المعقولية العظمى المعرفة للمقدرين (ML وPseudo-ML) هي نفسها، فلوغاريتم دالة المعقولية العظمى، الموافقة لعينة متكونة من T مشاهدة $(Y_1, Y_2, ..., Y_T)$ من Y_t، تحت فرضية القانون الشرطي الطبيعي لـ Y_t. تكتب من الشكل:

$$\log L(\theta) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log h_t(\theta) - \frac{1}{2}\sum_{t=1}^{T}\frac{[Y_t - m_t(\theta)]^2}{h_t(\theta)}$$

حيث $h_t(\theta)$ تمثل التباين الشرطي.

بتطبيق هذه الصيغة في حالة نموذج انحدار خطي بسيط ذي خطأ ARCH:

$$
\begin{aligned}
&Y_t = \beta X_t + \varepsilon_t \\
&\varepsilon_t = \eta_t h_t^{1/2}(\theta) \ , \ \ \eta_{\mathrm{t}} \sim N.i.d(0,1) \\
&E(\varepsilon_t \mid \varepsilon_{r-1}) = 0 \\
&\mathrm{var}(\varepsilon_t \mid \varepsilon_{r-1}) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{r-i}^2
\end{aligned}
$$

في هذه الحالة:

$$
\begin{aligned}
&E(Y_t \mid Y_{t-1}, X_t) = m_t(\theta) = \beta X_t \\
&\mathrm{var}(Y_t \mid Y_{t-1}, X_t) = h_t(\theta) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{r-i}^2
\end{aligned}
$$

مع:

$$\theta = (\beta, \alpha_0, \alpha_1, \ldots, \alpha_p) \in R^{q+2}$$

إذن لوغاريتم دالة المعقولية العظمى تكتب:

$$
\begin{aligned}
\log L(\theta) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log\left\{\alpha_0 + \sum_{i=1}^{p}\alpha_i (Y_{r-i} - \beta X_{t-i})^2\right\} \\
-\frac{1}{2}\sum_{t=1}^{T}(Y_r - \beta X_t)^2 \times \left[\alpha_0 + \sum_{i=1}^{p}\alpha_i (Y_{r-i} - \beta X_{t-i})^2\right]^{-1}
\end{aligned}
$$

إن المقدرات ML (أو Pseudo-ML) تحت فرضية التوزيع الطبيعي، نرمز لها $\hat{\theta}$ حيث $\theta \in R^k$ تحقق في مجملها نظام غير خطي يتكون من k معادلة:

$$\left.\frac{\partial \log L(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}} = 0$$

مع:

$$
\begin{aligned}
\left.\frac{\partial \log L(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}} = -\frac{1}{2}\sum_{t=1}^{T}\frac{1}{h_t(\hat{\theta})}\left.\frac{\partial h_t(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}} + \frac{1}{2}\sum_{t=1}^{T}\frac{\left[Y_t - m_t(\hat{\theta})\right]^2}{h_t^2(\hat{\theta})}\left.\frac{\partial h_t(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}} \\
+ \sum_{t=1}^{T}\frac{\left[Y_t - m_t(\hat{\theta})\right]}{h_t(\hat{\theta})}\left.\frac{\partial h_t(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}}
\end{aligned}
$$

إن هذا النظام يمكن أن يُقَسَّم إلى نظامين جزئيين، حسب المعالم θ الداخلة بشكل منفصل في صياغة المتوسط والتباين الشرطيين، كذلك إذا كان لدينا $\theta=(\alpha,\beta)'$، حيث α المعالم التابعة للمتوسط الشرطي، و β للتباين الشرطي، فإن:

$$\left.\frac{\partial \log L(\alpha)}{\partial \alpha}\right|_{\theta=\hat{\theta}} = \sum_{t=1}^{T}\left[\frac{Y_t - m_t(\hat{\alpha})}{h_t(\hat{\beta})}\right]\left.\frac{\partial m_t(\alpha)}{\partial \theta}\right|_{\alpha=\hat{\alpha}}$$

$$\left.\frac{\partial \log L(\beta)}{\partial \beta}\right|_{\theta=\hat{\theta}} = -\frac{1}{2}\sum_{t=1}^{T}\frac{1}{h_t(\hat{\beta})}\left.\frac{\partial h_t(\beta)}{\partial \beta}\right|_{\theta=\hat{\theta}} + \frac{1}{2}\sum_{t=1}^{T}\frac{[Y_t - m_t(\hat{\alpha})]^2}{h_t^2(\hat{\beta})}\left.\frac{\partial h_t(\beta)}{\partial \beta}\right|_{\beta=\hat{\beta}}$$

تحت عدة شروط وضوابط تعديليه نجد أن المقدر Pseudo-ML متقارب وطبيعي:

$$\sqrt{T}\left(\hat{\theta}-\theta\right)\xrightarrow[T\to\infty]{} N\left(0, J^{-1}IJ^{-1}\right)$$

أما مصفوفة التباين- التباين المشترك المتقاربة للمقدر Pseudo-ML فإنها تُحسّب من خلال:

$$J = E_0\left[-\frac{\partial^2 \log L(\theta)}{\partial\theta\partial\theta'}\right]$$

$$I = E_0\left[\frac{\partial \log L(\theta)}{\partial \theta}\frac{\partial \log L(\theta)}{\partial \theta'}\right]$$

حيث E_0 يمثل المتوسط المأخوذ حسب اختلاف القانون. في الحالة التطبيقية، المصفوفتان I و J تُقَدَّرَان مباشرة باستبدال المتوسط E_0 بالمتوسط التجريبي (أو التقديري) والمعلم غير المعروف θ بالمقدر المتقارب $\hat{\theta}$ "Asymptotic Estimator"، لدينا:

$$\hat{J} = -\frac{1}{T}\sum_{t=1}^{T}\left.\frac{\partial^2 \log L(\theta)}{\partial\theta\partial\theta'}\right|_{\theta=\hat{\theta}}$$

$$\hat{I} = \frac{1}{T}\sum_{t=1}^{T}\left.\frac{\partial \log L(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}}\left.\frac{\partial \log L(\theta)}{\partial \theta'}\right|_{\theta=\hat{\theta}}$$

و التباين المقدر لـ . $\hat{\theta}$ يحقق:

$$\operatorname{var}\left[\sqrt{T}\left(\hat{\theta}-\theta\right)\right] = \hat{J}^{-1}\hat{I}\hat{J}^{-1}$$

في حالة $J = I$ (Maximum Likelihood) تصبح مصفوفة التباين- التباين المشترك المُقَارِبَة من الشكل:

$$\text{var}\left[\sqrt{T}\left(\hat{\theta}-\theta\right)\right]=J^{-1}$$

في حالة ML لما يكون بالإمكان فصل معالم المتوسط الشرطي والتباين الشرطي، نستطيع أن نبين:

$$\text{var}\left[\sqrt{T}\left(\hat{\theta}-\theta\right)\right]=\left[\frac{1}{T}\sum_{t=1}^{T}\frac{1}{2h_t^2(\hat{\beta})}\frac{\partial h_t(\beta)}{\partial\beta}\bigg|_{\beta=\hat{\beta}}\frac{\partial h_t(\beta)}{\partial\beta'}\bigg|_{\beta=\hat{\beta}}\right]^{-1}$$

يمكن أيضا تطبيق طريقة المربعات الصغرى المعممة لتقدير معالم النموذج ARCH. نتبع الخطوات التالية:

الخطوة الأولى: تقدير نموذج الانحدار الكلاسيكي $Y = X\beta + \varepsilon$

الخطوة الثانية: انطلاقا من بواقي التقدير $\hat{\varepsilon}_t$، نقدر الانحدار $\hat{\varepsilon}_t^2 = \alpha_0 + \sum_{i=1}^{p}\alpha_i\hat{\varepsilon}_{t-i}^2 + \eta_t$ بطريقة المربعات الصغرى العادية، فنتحصل على المعالم المقدرة $\left(\hat{\alpha}_0, \hat{\alpha}_1, ..., \hat{\alpha}_p\right)$

الخطوة الثالثة: نقوم بحساب التباين الشرطي انطلاقا من المعادلة $\hat{h}_t = \hat{\alpha}_0 + \sum_{i=1}^{p}\hat{\alpha}_i\hat{\varepsilon}_{t-i}^2$

الخطوة الرابعة: نعيد تقدير شعاع المعالم β بطريقة المربعات الصغرى العادية في النموذج الجديد:

$$\frac{Y_t}{\sqrt{\hat{h}_h}}=\frac{\beta_0}{\sqrt{\hat{h}_h}}+\beta_1\frac{X_{t1}}{\sqrt{\hat{h}_h}}+\beta_2\frac{X_{t2}}{\sqrt{\hat{h}_h}}+...+\beta_k\frac{X_{tk}}{\sqrt{\hat{h}_h}}+\frac{\varepsilon_t}{\sqrt{\hat{h}_h}}$$

وهذا يعني أنه يتم إعادة تقدير المعالم بطريقة GLS:

$$\tilde{\beta}=(X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}Y)$$

أي يتعلق الأمر بانحدار مرجح (ذات أوزان) مع معامل الترجيح $\omega=\frac{1}{\sqrt{\hat{h}_t}}$ و $\Omega = diag(h_t)$ وبعد ذلك، يمكن تحسين تقدير المعاملات α_i بطريقة GLS أي $\tilde{\alpha}=(\tilde{\varepsilon}'\Omega^{-1}\tilde{\varepsilon})^{-1}(\tilde{\varepsilon}'\Omega^{-1}h^*)$.

نشير هنا إلى أن تباين الأخطاء غير ثابت و مجال الثقة للتنبؤ دالة لتطاير السلسلة في حد ذاتها. يمكن القول أن هذه الطريقة تعتبر سهلة الاستخدام ولكن أقل فعالية من طريقة المعقولية العظمى.

أشار Weiss (1986) إلى إمكانية إدخال على التباين الشرطي تأثيرات إضافية Additional Effects للمتغير المُفَسَّر حيث أن من خواص نمذجة GARCH أنها تسمح بإضافة هذه القوى سواءً من خلال المتوسط الشرطي، أو من خلال التباين الشرطي، فمثلا يمكن لنا أن نتصور نموذج ARMA حيث يكون للتباين غير الشرطي لـ Y. تأثير على التباين الشرطي:

$$\phi(L)Y_t = \theta(L)\varepsilon_t$$
$$E(\varepsilon_t \mid I_{r-1}) = 0$$
$$\varepsilon_t = \eta_t \times h_t^{1/2} \quad , \quad \eta_t \sim N(0,1)$$
$$h_t = \mathrm{var}(\varepsilon_t \mid I_{t-1}) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \gamma_0 \left[E(Y_t \mid Y_{t-1})\right]^2 + \sum_{i=1}^{p} \gamma_i v_{t-i}^2$$

وهذا يعني أنه يمكن دراسة نموذج ARMA بخطأ GARCH، حيث يكتب النموذج ARMA-GARCH كما يلي:

$$\phi(L)Y_t = \theta(L)\varepsilon_t$$
$$\varepsilon_t \mid \varepsilon_{t-1} \sim N(0, h_t)$$
$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}$$
$$\alpha_0 > 0, \quad \alpha_i \geq 0, \quad i = 1,\dots, p, \ \beta_j \geq 0,$$

$$j = 1,\dots, q$$

حسب ما أشار إليه Gouriéroux (1992) فإن الطرق الممكنة لتقدير التباين الشرطي ترتكز على اقتراح مجالات ثقة للمتغير المُفَسَّر (Endogenous Variable) مبنية على عدم وضع صفة الثبات مع الزمن للعزوم من الرتبة (الدرجة) 2. لهذا يمكن القول أن الفرق الأساسي بين نمذجة ARMA وARCH يكمن في أن مجال الثقة للأولى مبني على تباين ثابت مع الزمن، وهذا مالا نجده في نموذج ممثل بـ . GARCH-ARCH للبواقي.

هناك طريقتان مختلفتان لتحليل هذا النموذج[1]. تتمثل الأولى في الطرق الكلاسيكية في تقدير وتحليل السيرورة ARMA، أي كما لو أن لدينا معطيات ذات تباين شرطي غير متجانس للأخطاء وتكون هنا مقدرات معاملات كثيرات الحدود ϕ ، θ متقاربة (convergent). في هذه الحالة، التنبؤ بأفق واحد لـ . Y_t، ونعني به المتغيرات $\hat{Y}_t = \left[\frac{\hat{\phi}(L)}{\hat{\theta}(L)} - 1\right] Y_t$، التي تكون تحت شروط تعديلية غير متحيزة، حيث $\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \hat{Y}_t\right)^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t^2$. أما مجالات التنبؤ فهي: $\left[\hat{Y}_t \pm 2\hat{\sigma}\right]$ (نهمل أثر مقدرات ϕ و θ). $\hat{\sigma}^2$ هو مقدر متقارب لـ . $E\left(h_t\right)^2$. أي القيمة المتوسطة لسرعة التقلبات، وهي حالة خاصة مستقلة عن الفترة t للتنبؤ، لما يكون لكل مجالات التنبؤ نفس الطول. أما الثانية، فيمكن الأخذ بعين الاعتبار نموذج تطور سرعة التقلبات وتطبيق خطوات التقدير المخصصة لنماذج ARCH. إذا كانت $\hat{\hat{\theta}}, \hat{\hat{\phi}}$ تمثلان مقدري نموذجي الانحدار الذاتي والمتوسطات المتحركة (على الترتيب)، فإن التنبؤ بأفق واحد لـ . Y_t يعطى كما يلي:

$$\hat{\hat{Y}}_t = \left[\frac{\hat{\hat{\phi}}(L)}{\hat{\hat{\theta}}(L)} - 1\right] Y_t$$

تكون هذا الأخيرة تحت شروط تعديلية غير متحيزة. وفي هذه الحالة مجالات التنبؤ تحسب من العلاقة $\left[\hat{\hat{Y}}_t \pm z_{\alpha/2}\hat{h}_t\right]$، حيث $\hat{h}_t$ مقدر سرعة التقلبات (التباين الشرطي) في اللحظة t. إذن طول مجالات التنبؤ هنا مرتبط بالزمن t.

1- Droesbeke, Fichet and Tassi, p. 82.

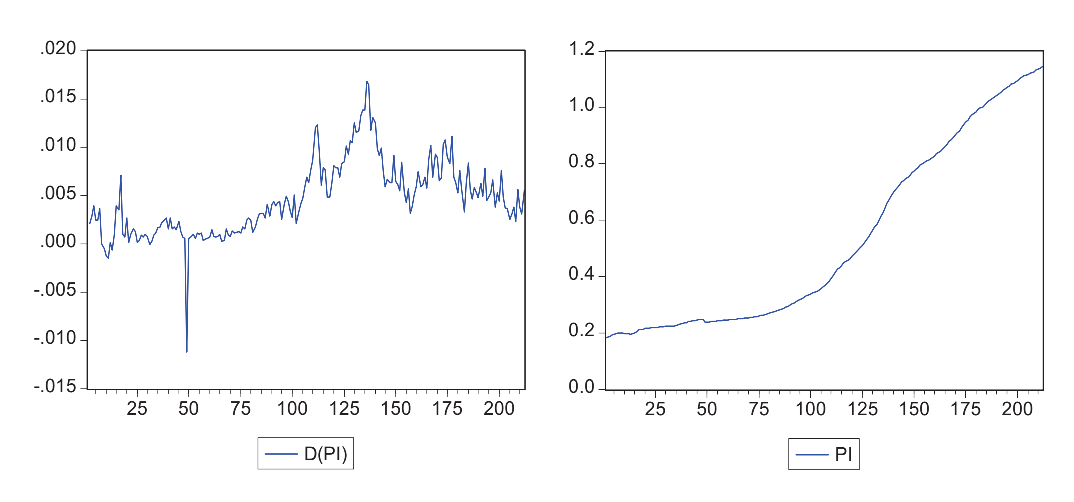
2- لدينا:

$$E\left[Y_t - E\left(Y_t \mid Y_{t-1}\right)\right]^2 = E\left\{E\left[\left[Y_t - E\left(Y_t \mid \mathrm{Y}_{t\text{-}1}\right)\right]^2 \mid Y_{t-1}\right]\right\} = E\ \left\{\mathrm{var}\left(Y_t \mid \mathrm{Y}_{t\text{-}1}\right)\right\} = E\left(h_t\right)$$

مثال 2:

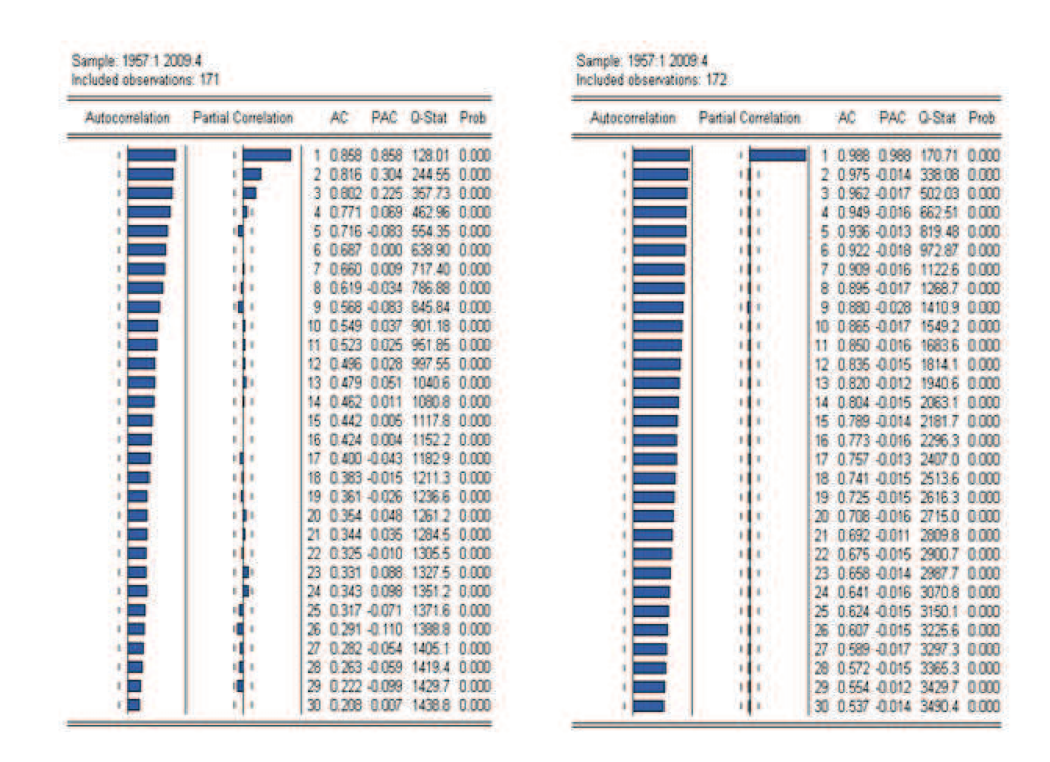
لتكن سلسلة الرقم القياسي للأسعار في الولايات المتحدة الأمريكية التي تحتوي على 212 مشاهدة. في هذا المثال، سنستخدم برمجيتي RATS 5.04 و Eviews 5.0. الشكل البياني التالي يظهر تطور هذه السلسلة:

الشكل رقم (4): التمثيل البياني لمستوى الرقم القياسي للأسعار
(الأصلية و ذات الفروقات من الدرجة الأولى)



نقوم أولا بدراسة استقرارية السلسلة. من خلال الشكل (5)، يتبين لنا أن هذه السلسلة PI غير مستقرة لأن معاملات الارتباط الذاتي كلها تختلف معنويا عن الصفر بنسبة معنوية 0.05 أي أنها تقع كلها خارج مجال الثقة $\left[-\frac{1.96}{\sqrt{T}},+\frac{1.96}{\sqrt{T}}\right]$ مما يقودنا إلى حساب الفروقات من الدرجة الأولى D(PI) (أنظر أيضا الشكل (4).

الشكل رقم (5): التمثيل البياني لدالة الارتباط الذاتي والجزئي للسلسلة (الأصلية و ذات الفروقات من الدرجة الأولى (على الترتيب)



Sample: 1957:1 2009:4
Included observations: 172

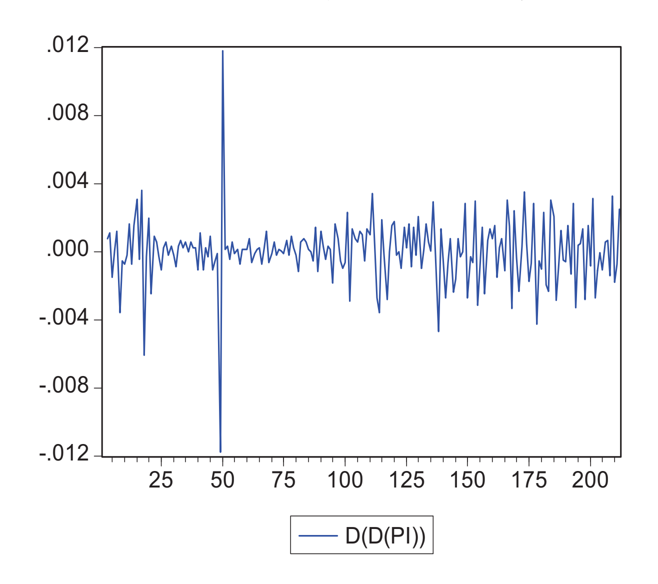
	AC	PAC	Q-Stat	Prob
1	0.988	0.988	170.71	0.000
2	0.975	-0.014	338.08	0.000
3	0.962	-0.017	502.03	0.000
4	0.949	-0.016	662.51	0.000
5	0.936	-0.013	819.48	0.000
6	0.922	-0.018	972.87	0.000
7	0.909	-0.016	1122.6	0.000
8	0.895	-0.017	1268.7	0.000
9	0.880	-0.028	1410.9	0.000
10	0.865	-0.017	1549.2	0.000
11	0.850	-0.016	1683.6	0.000
12	0.835	-0.015	1814.1	0.000
13	0.820	-0.012	1940.6	0.000
14	0.804	-0.015	2063.1	0.000
15	0.789	-0.014	2181.7	0.000
16	0.773	-0.016	2296.3	0.000
17	0.757	-0.013	2407.0	0.000
18	0.741	-0.015	2513.6	0.000
19	0.725	-0.015	2616.3	0.000
20	0.708	-0.016	2715.0	0.000
21	0.692	-0.011	2809.8	0.000
22	0.675	-0.015	2900.7	0.000
23	0.658	-0.014	2987.7	0.000
24	0.641	-0.016	3070.8	0.000
25	0.624	-0.015	3150.1	0.000
26	0.607	-0.015	3225.6	0.000
27	0.589	-0.017	3297.3	0.000
28	0.572	-0.015	3365.3	0.000
29	0.554	-0.012	3429.7	0.000
30	0.537	-0.014	3490.4	0.000

Sample: 1957:1 2009:4
Included observations: 171

	AC	PAC	Q-Stat	Prob
1	0.858	0.858	128.01	0.000
2	0.816	0.304	244.55	0.000
3	0.802	0.225	357.73	0.000
4	0.771	0.069	462.96	0.000
5	0.716	-0.083	554.35	0.000
6	0.687	0.000	638.90	0.000
7	0.660	0.009	717.40	0.000
8	0.619	-0.034	786.88	0.000
9	0.568	-0.083	845.84	0.000
10	0.549	0.037	901.18	0.000
11	0.523	0.025	951.85	0.000
12	0.496	0.028	997.55	0.000
13	0.479	0.051	1040.6	0.000
14	0.462	0.011	1080.8	0.000
15	0.442	0.006	1117.8	0.000
16	0.424	0.004	1152.2	0.000
17	0.400	-0.043	1182.9	0.000
18	0.383	-0.015	1211.3	0.000
19	0.361	-0.026	1236.6	0.000
20	0.354	0.048	1261.2	0.000
21	0.344	0.036	1284.5	0.000
22	0.325	-0.010	1305.5	0.000
23	0.331	0.088	1327.5	0.000
24	0.343	0.098	1351.2	0.000
25	0.317	-0.071	1371.6	0.000
26	0.291	-0.110	1388.8	0.000
27	0.282	-0.054	1405.1	0.000
28	0.263	-0.059	1419.4	0.000
29	0.222	-0.099	1429.7	0.000
30	0.208	0.007	1438.8	0.000

من الملاحظ أن السلسلة الجديدة لا تتذبذب حول وسط حسابي ثابت أي ما زالت تحتوي على مركبة الاتجاه العام. ففي هذه الحالة، لا بد من تحويل السلسلة عن طريق حساب الفروقات من الدرجة الثانية D(DPI)) (أنظر الشكل (6)) أي $d = 2$.
الشكل (7) يظهر تطور دالة الارتباط الذاتي و الجزئي للسلسلة الجديدة:

الشكل رقم (6): التمثيل البياني للسلسلة المحولة D(D(PI)



الشكل رقم (7): التمثيل البياني لدالة الارتباط الذاتي و الجزئي

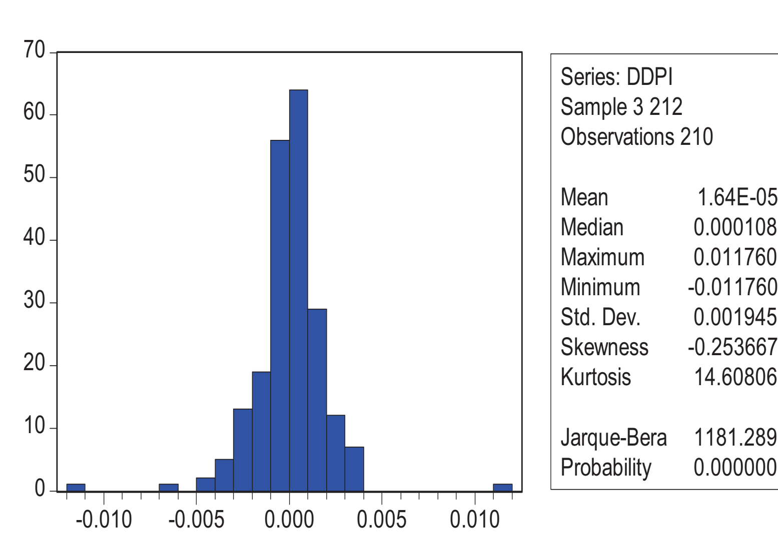
Sample: 1 212
Included observations: 210

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.354	-0.354	26.708	0.000
		2	-0.088	-0.244	28.360	0.000
		3	0.081	-0.054	29.758	0.000
		4	0.026	0.025	29.908	0.000
		5	-0.067	-0.035	30.877	0.000
		6	-0.031	-0.075	31.083	0.000
		7	-0.010	-0.088	31.106	0.000
		8	0.068	0.022	32.113	0.000
		9	-0.088	-0.063	33.844	0.000
		10	0.029	-0.016	34.036	0.000
		11	-0.001	-0.032	34.036	0.000
		12	-0.047	-0.076	34.526	0.001
		13	0.044	-0.004	34.970	0.001
		14	-0.019	-0.025	35.054	0.001
		15	-0.015	-0.029	35.103	0.002
		16	0.018	-0.017	35.178	0.004
		17	-0.015	-0.029	35.232	0.006
		18	0.034	0.015	35.496	0.008
		19	-0.057	-0.053	36.260	0.010
		20	-0.001	-0.048	36.260	0.014
		21	0.041	-0.009	36.648	0.018
		22	-0.080	-0.085	38.181	0.017
		23	-0.016	-0.093	38.242	0.024
		24	0.140	0.080	42.938	0.010
		25	-0.018	0.078	43.019	0.014
		26	-0.057	-0.004	43.801	0.016
		27	0.012	-0.027	43.838	0.021
		28	0.122	0.101	47.453	0.012
		29	-0.146	-0.065	52.675	0.005
		30	0.054	0.033	53.387	0.005

من خلال الشكل البياني الممثل أعلاه، يظهر جليا أن معاملات الارتباط الذاتي للسلسلة D(D(PI)) تساوي معنويا الصفر، أي تقع داخل مجال الثقة، نستنتج من ذلك أن السلسلة مستقرة (درجة التفاضل تساوي 2 أي $PI \sim I(2)$).

إن السلسلة المستقرة D(D(PI)) تخضع لتوزيع غير طبيعي (إحصائية Jarque-Bera أكبر تماما من القيمة الحرجة لتوزيع χ^2) ومعامل Skewness في هذه الحالة سالب (ملتو نحو اليسار)، يدل على عدم تناظر التوزيع الاحتمالي ومعامل Kurtosis أكبر تماما من 3 (أنظر الشكل رقم (8))، فعدم التناظر يمكن أن يكون بسبب وجود بنية غير خطية في سلسلة الرقم القياسي للأسعار، فمثلا يمكن أن يكون السبب في وجود تباين شرطي غير متجانس (تأثير ARCH).

الشكل رقم (8): اختبار التوزيع الطبيعي



بالنظر إلى الشكل (7)، نلاحظ أن معامل الارتباط $\rho(1)$ يختلف معنويا عن الصفر (أي يقع خارج مجال الثقة) ومن أجل $k>1$ كل معاملات الارتباط الذاتي تنعدم معنويا، وهي الحالة التي توافق نموذج MA(1)، كما نلاحظ أيضا أن معاملي الارتباط الجزئي $r(1)$ و $r(2)$ يختلفان معنويا عن الصفر ومن أجل $k>1$ كل معاملات الارتباط الجزئي تنعدم معنويا، وهي الحالة التي توافق نموذج AR(2).

وفقا لهذه النقاط تكون الصيغة الرياضية المثلى للنموذجين المرشحين المعرفين للسلسلة المستقرة من الشكل:

$$ARIMA(0,2,1): \nabla Y_t = \delta + (1+\theta_1 L)\varepsilon_t$$
$$ARIMA(2,2,0): (1-\phi_1 L)\nabla Y_t = \delta + \varepsilon_t$$

وبعد تقدير هذين النموذجين، يكون النموذج المختار هو الذي يُعطي أحسن توفيقة بين المعايير Akaike، Schwarz، أي تصغير هذين المعيارين.

الجدول (1): اختيار النموذج الأمثل

معيار AIC / Schwarz	نوع النموذج المرشح
-9.814 / -9.766	$ARIMA(0,2,1)$
-9.826 / -9.794	$ARIMA(2,2,0)$

نلاحظ أن النموذج الأمثل الذي يعبر أكثر عن تغيرات الرقم القياسي للأسعار هو نموذج(ARIMA(2,2,0. نتائج التقدير تظهر باستعمال برمجية RATS:

```
boxjenk(noprint,ar=||1,2||,diff=2,span=4,ma=0) y / resids

Box-Jenkins - Estimation by Gauss-Newton
Convergence in     2 Iterations. Final criterion was  0.0000000 <  0.0000100
Dependent Variable Y
Quarterly Data From 1948:01 To 1999:04
Usable Observations    208     Degrees of Freedom   206
Centered R**2      0.999971      R Bar **2   0.999970
Uncentered R**2    0.999992      T x R**2     207.998
Mean of Dependent Variable      0.5350945529
Std Error of Dependent Variable 0.3260929534
Standard Error of Estimate      0.0017714179
Sum of Squared Residuals        0.0006464118
Durbin-Watson Statistic             2.016620
Q(36-2)                            22.573872
Significance Level of Q           0.93285454

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  AR{1}                  -0.445685680  0.067680563     -6.58514  0.00000000
2.  AR{2}                  -0.246697565  0.067679202     -3.64510  0.00033842
```

نلاحظ أن للمعالم معنوية إحصائية بنسبة معنوية 0.05 باعتبار أن قيم ستيودنت بالقيمة المطلقة أكبر تماما من القيمة الحرجة للتوزيع الطبيعي، إضافة إلى ذلك، للنموذج قدرة تفسيرية عالية جدا. من خلال الشكل (9)، نستنتج أن سلسلة البواقي مستقرة حيث أن معاملات الارتباط الذاتي تقع كلها داخل مجال الثقة $\left[\frac{-1.96}{\sqrt{T}},\frac{+1.96}{\sqrt{T}}\right]$ و هذا يعني أن هناك استقلالية تامة بين الأخطاء، إلا أن اختبار ARCH(1) يظهر عدم تجانس التباين الشرطي للأخطاء حيث أن إحصائية ARCH-LM ($LM = 12.516$) أكبر تماما من القيمة المجدولة لتوزيع χ^2 بدرجة حرية 1، نرفض إذن فرضية تجانس التباين الشرطي H_0 وهذا يعني أن البواقي تخضع لنموذج GARCH الذي ينبغي أن نختبر معنويته الإحصائية، لتكن النتائج التالية باستخدام RATS:

```
set sqr_lag_resids / = resids{1}**2
set sqr_resids / = resids**2
linreg(noprint) sqr_resids / u
#constant sqr_lag_resids
compute arch_lm_stat=%nobs*%rsquared
display arch_lm_stat
      12.51680
```

وكنتيجة لذلك، يتم تقدير النموذج ARIMA(2,2,0)-GARCH(1,1) باستخدام خوارزمية BHHH (وهي إحدى الطرق التي تستخدم في النماذج غير الخطية بتعظيم لوغاريتم دالة المعقولية)، فنحصل على النتائج التالية المبينة في الجدول (2):

الجدول (2): تقدير النموذج ARIMA(2,2,0)-GARCH(1,1)

إحصائية ستيودنت	القيم المقدرة	المعالم
-4.9703	-0.3858	$\hat{\phi}_1$
-3.2432	-0.2142	$\hat{\phi}_2$
2.6353	0.1267	$\hat{\alpha}_1$
2.5721	0.2172	$\hat{\beta}_1$

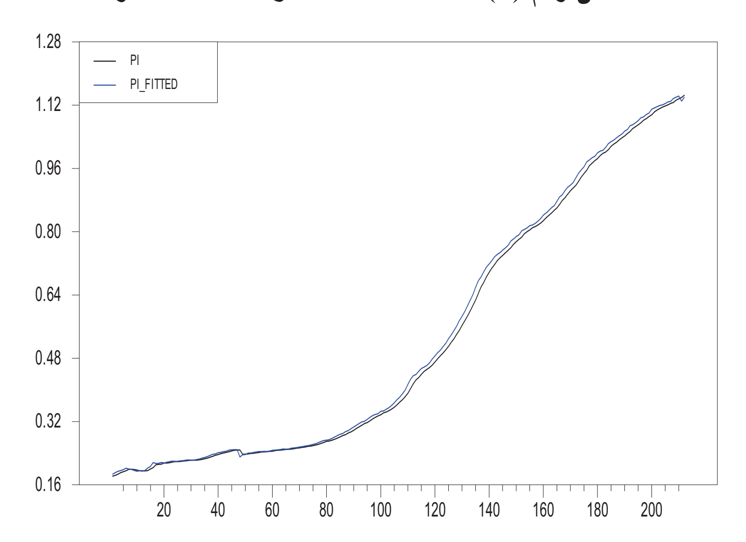
المعالم	القيم المقدرة	إحصائية ستيودنت
$\hat{\alpha}_0$	0.0000019	587.8510
R^2	0.9999	
مجموع مربعات البواقي $\sum \hat{\varepsilon}_t^2$	0.00064	
إحصائية دربين–واتسون	2.0371	
دالة المعقولية المقدرة	1041.2977	
AIC / Schwarz	-9.8599 / -9.7643	
إحصائية White	0.0014	
إحصائية ARCH-LM	0.1014	

من خلال النتائج المتحصل عليها من الجدول (2)، يمكن قبول النموذج المقترح وذلك للاعتبارات التالية:

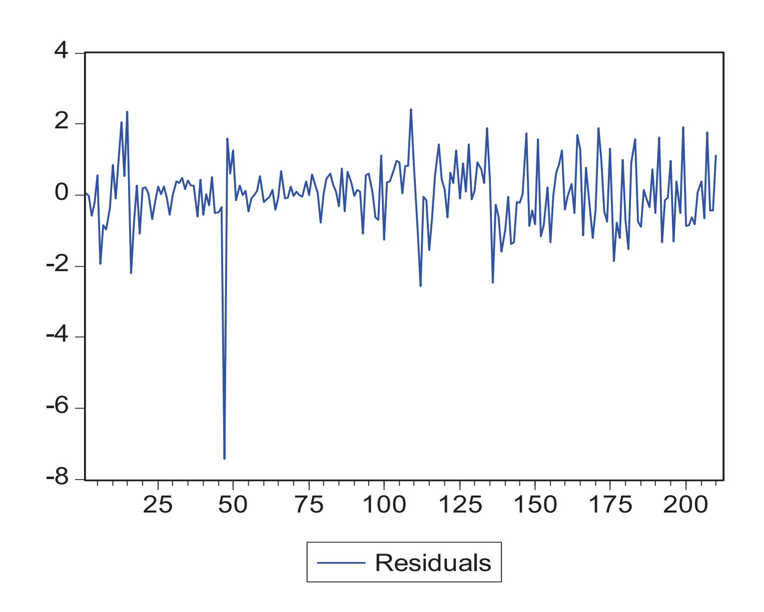
- لجميع المعالم معنوية إحصائية، أي أنها تختلف معنويا عن الصفر بنسبة معنوية 0.05. نرفض H_0 (قيم ستيودنت أكبر تماما من القيمة الحرجة للتوزيع الطبيعي 1.96)،
- معاملا نموذج GARCH المقدران $\hat{\alpha}_1$ و $\hat{\beta}_1$ موجبان و أيضا $\hat{\alpha}_1 + \hat{\beta}_1 = 0.1267 + 0.2172 = 0.3439 < 1$ أي أن شرط الاستقرارية محقق.
- للنموذج ARIMA-GARCH قدرة تفسيرية عالية جدا ($R^2 = 0.999$)، وهذا ما نلاحظه من خلال تطابق السلسلة الأصلية مع تلك المقدرة (أنظر الشكل (9))،
- هناك استقلالية تامة بين بواقي التقدير من خلال نتيجة إحصائية دربين–واتسون التي تساوي 2.0371،
- تشير إحصائية White إلى أن التباين الهامشي لأخطاء النموذج ARIMA-GARCH متجانس، حيث أن إحصائية LM أقل تماما من القيمة المجدولة لتوزيع χ^2 بدرجة حرية 8،
- التباين الشرطي لهذه الأخطاء متجانس باعتبار أن إحصائية ARCH-LM التي تساوي 0.1014 أقل تماما من القيمة المجدولة لتوزيع χ^2 بدرجة حرية 1. يمكن

التأكد من ذلك من خلال دالة الارتباط الذاتي لمربعات البواقي (أنظر الشكل (11)). نلاحظ أن سلسلة مربعات البواقي مستقرة لأن كل معاملات الارتباط الذاتي تساوي معنويا الصفر أي تقع كلها داخل مجال الثقة.

الشكل رقم (9): السلسلة الأصلية و السلسلة المقدرة



الشكل رقم (10): بواقي تقدير نموذج ARIMA(2,2,0)-GARCH(1,1)



الشكل رقم (11): دالة الارتباط الذاتي و الجزئي لمربعات البواقي

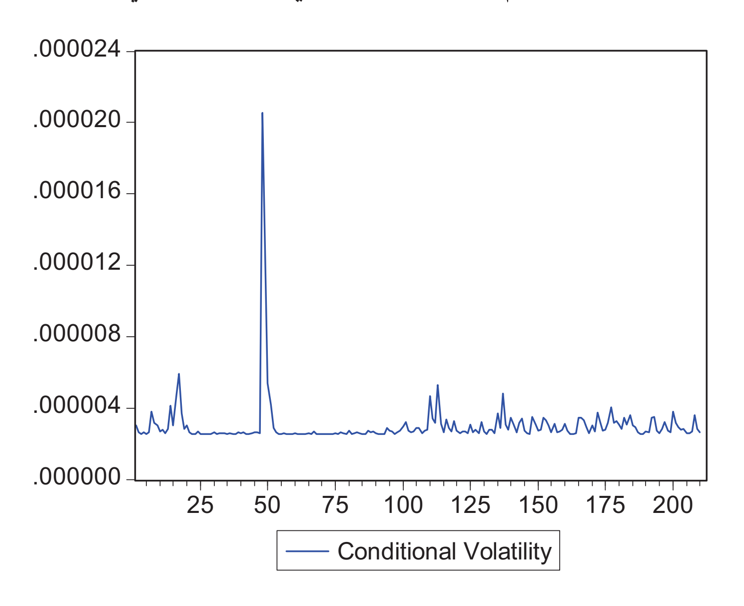
Sample: 1 212
Included observations: 210

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.022	0.022	0.1043	0.747
		2	-0.011	-0.011	0.1290	0.938
		3	0.023	0.023	0.2393	0.971
		4	-0.020	-0.021	0.3241	0.988
		5	-0.025	-0.023	0.4559	0.994
		6	-0.022	-0.022	0.5659	0.997
		7	-0.016	-0.014	0.6188	0.999
		8	-0.013	-0.012	0.6570	1.000
		9	-0.019	-0.018	0.7344	1.000
		10	-0.026	-0.026	0.8812	1.000
		11	-0.028	-0.028	1.0512	1.000
		12	-0.014	-0.015	1.0983	1.000
		13	-0.029	-0.030	1.2888	1.000
		14	-0.030	-0.031	1.4887	1.000
		15	-0.026	-0.028	1.6383	1.000
		16	-0.015	-0.018	1.6916	1.000
		17	-0.025	-0.029	1.8339	1.000
		18	-0.024	-0.028	1.9685	1.000
		19	-0.012	-0.018	2.0044	1.000
		20	-0.031	-0.038	2.2302	1.000
		21	-0.028	-0.034	2.4103	1.000
		22	-0.027	-0.035	2.5816	1.000
		23	-0.028	-0.036	2.7643	1.000
		24	-0.005	-0.015	2.7703	1.000
		25	-0.023	-0.034	2.8928	1.000
		26	-0.033	-0.045	3.1563	1.000
		27	-0.014	-0.028	3.2038	1.000
		28	-0.004	-0.019	3.2082	1.000
		29	-0.021	-0.036	3.3161	1.000
		30	-0.011	-0.028	3.3485	1.000

يمكن تمثيل تغيرات التباين الشرطي بيانيا:



الشكل رقم (12): التمثيل البياني للتباين الشرطي



لقد بينت دراسات كثيرة أن نمذجة ARCH أو GARCH تفترض وجود علاقة تربيعية بين الخطأ و التباين الشرطي. يمكن الأخذ بهذه الصيغة في حالة ما إذا كانت للظواهر التي يتم تحليلها نفس الإشارة، فاختيار الشكل التربيعي للتباين الشرطي له نتائج مهمة على المسار الزمني للسلسلة.

توجد علاقة سالبة بين المردودية الحالية و المخطر المستقبلي، و هذا ما يطرح مشكل عدم الخطية. انخفاض في قيمة الأسهم يؤدي إلى ارتفاع نسبة الاستدانة متبوعا أيضا بارتفاع في مخطر الإفلاس الذي يحدث بسبب تزايد في سرعة التقلبات المستقبلية، نجده يرتبط ارتباطا سالبا مع المردودية الحالية للسهم. هذا النوع من النماذج لا يمكن أن يتلاءم مع حركية هذه الظواهر المالية لأن التباين الشرطي في هذه الحالة لا يعتمد إلا على التباينات ومربعات ا؟لأخطاء في الفترات السابقة، ليس لإشارة المردودية أي تأثير على التقلبات Volatility. إضافة إلى ذلك، القرارات المتخذة من طرف الوكلاء الاقتصاديين تعتمد على عدة متغيرات و نمذجة ARCH لا تهتم إلا بسلوك فردي لعدة أصول مالية.

4. النماذج المستحدثة عن الانحدار الذاتي ذات التباين الشرطي غير المتجانس

يوجد عدة نماذج ظهرت كنتيجة للانتقادات التي وجهت إلى نموذج ARCH. ك ل نموذج يبحث في الأخذ بعين الاعتبار حدث معين في السلاسل المالية أو يسمح بدراس ة الظواهر الاقتصادية الكلية (قياس عدم التأكد في مسألة التضخم، دراسة آثار ت دخلات البنك المركزي،...الخ). هذه النماذج لا تأخذ بعين الاعتبار فقط مدى بواقي المتوس ط الشرطي، بل أيضا إشاراتها بإدخال آليات عدم التناظر. نذكر منها:

1.4. نماذج ARCH / GARCH غ ير المتن اظرة Asymmetric ARCH or GARCH Models:

إن من أهم المقاربات التي تغطي النماذج ARCH غير الخطية تلك التي تأخذ في الحسبان الظواهر غير المتناظرة، وترتكز على فكرة بسيطة هي أن مفعول (تأثير) عدم تجانس التباين يختلف هنا حسب كون إشارة الخطأ السابق (موجبة أو سالبة)، حيث نجد مجموعتين من هذه النماذج: نموذج EGARCH Exponential Generalized AutoRegressive Conditional Heteroskedastic التي اقترحها Nelson (1991) حيث اهتم بالتطور غير المتناظر للتباين ونماذج TARCH " Threshold AutoRegressive Conditional heteroscedastic" أو ما يعرف بنماذج ARCH ذات العتبة التي اقترحها Engle and Bollerslev (1986) وفي سنة 1991 عَمَّمَ Rabemananjara and Zakoian هذه النماذج وأطلقا على تسميتها بنموذج TGARCH " Threshold Generalized AutoRegressive Conditional Heteroscedastic".

نقول أن السيرورة ε_t توافق نموذج EARCH(p) إذا وفقط إذا كان:

$$\log h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i [\phi \eta_{t-i} + \gamma(|\eta_{t-i}| - E(|\eta_{t-i}|))]$$

$$\eta_{t-i} = \varepsilon_{t-i} / h_{t-i}$$

$$E(\eta_t \mid I_t) = 0$$

الحد غير المتناظر الذي يدخل في هذا النموذج يتمثل في الدالة $g(\eta_t) = \phi \eta_{t-i} + \gamma(|\eta_{t-i}| - E(|\eta_{t-i}|))$ التي تعبر عن سلسلة ذات توزيع متماثل و مستقل

iid بمتوسط معدوم. الميزة الأساسية في هذا النموذج تتمثل في عدم وجود قيود عدم المساواة. باعتبار أن صياغة التباين الشرطي هي باللوغاريتم، المعالم α_i قد تكون موجبة أو سالبة.

أما السيرورة EGARCH(p) تكتب كما يلي:

$$\log h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i [\phi \eta_{t-i} + \gamma(|\eta_{t-i}| - E(|\eta_{t-i}|))] + \sum_{j=1}^{q} \beta_j \log h_{t-j}$$

يشير Nelson (1991) إلى أن اختيار المتغيرات η_{t-i} عوضا عن ε_{t-i} يسمح بالحصول على شروط استقرارية ضعيفة تتعلق فقط بكثير الحدود $\beta(L)$.

يوجد وجه أخر من النماذج غير المتناظرة، تسمى بنموذج TARCH ونموذج TGARCH المقترحة من طرف Zakoian (1991,1994) حيث الصيغة التربيعية تعوض بدالة خطية بـ . "قطعة" — كل قسم يرتبط بصدمات لها نفس الطبيعة — مما يسمح بالحصول على دوال مختلفة للتطاير الشرطي حسب الإشارة وقيم الصدمات. نقول أن السيرورة ε_t تحقق النموذج TGARCH(p,q) إذا وفقط إذا كان:

$$h_t^{1/2} = \alpha_0 + \sum_{i=1}^{p} \alpha_i^+ \varepsilon_{t-i}^+ - \sum_{i=1}^{p} \alpha_i^- \varepsilon_{t-i}^- + \sum_{j=1}^{q} \beta_j h_{t-j}$$

حيث $\varepsilon_t^+ = \max(\varepsilon,0)$ و $\varepsilon_t^- = \min(\varepsilon,0)$ مع $\alpha_0 > 0,$ ، $\alpha_i^+ \geq 0,$ ، $\alpha_i^- \geq 0$ و $\beta_j \geq 0,$ ، $\forall i,$ ، $\forall j$.

ترتكز هذه العبارة على نمذجة الانحراف المعياري الشرطي. يمكن إزالة قيود الإشارة على المعالم مما يسمح بالأخذ بعين الاعتبار ظواهر عدم التناظر، أثر صدمة ε_{t-i} على التباين الشرطي يرتبط بمدى و إشارة هذه الصدمة.

2.4. نماذج GARCH-M و GARCH-DLM:

إن تقييم المخطر يعتبر نقطة مهمة في الاقتصاد المالي. فلقد أشار Engle, Lilien and Robins (1987) إلى أن طرق قياس المخطر و التنبؤ به غالبا ما تكون بسيطة و بالتالي فهي غير مناسبة لتحليل السلاسل المالية. من الواضح أن التعويض الذي يتحصل عليه الوكلاء

نتيجة امتلاكهم للأسهم يجب أيضا أن يتغير بتغير الفترات الزمنية، فلا بد إذن من الأخذ بالحسبان تغيرات المخطر بدلالة الزمن. نمذجة ARCH-M "ARCH-in-Mean" تأخذ بعين الاعتبار هذه الظاهرة بإدخال التباين الشرطي في معادلة التوقع الشرطي، كل تغير في التباين الشرطي ينتج عنه تغير في المردودية المتوقعة للمحفظة. في هذه الحالة، التوقع الشرطي يرتبط بالتباين الشرطي الذي يصبح متغيرا مفسرا (مستقلا):

$$Y_t = X\beta + f(h_t) + \varepsilon_t$$

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2$$

$$\varepsilon_t = \eta_t \times h_t^{1/2}$$

في الواقع، التغيرات في التباين الشرطي مصحوب بتغير في التوقع الشرطي و عليه يمكن كتابة صيغة GARCH-M "GARCH-in-Mean" كما يلي:

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}$$

اقترح Cocco and Paruolo (1990) نموذجا يأخذ بعين الاعتبار التزايد في التقلبات (الفروقات من الدرجة الأولى) الذي يؤثر على المتغير التابع (الظاهرة المالية) يسمى هذا النوع من النماذج بـ GARCH-DM "Difference in Mean ":

$$Y = X\beta + f(h_t - h_{t-1}) + \varepsilon_t$$

يمكن أيضا ادخال الآثار الماضية بمساعدة دالة كثير الحدود للتباطؤات. يسمى هذا النموذج بـ GARCH-DLM "Distributed Lag in Mean" :

$$Y_t = X\beta + \nabla(L)(h_t - h_{t-1}) + \varepsilon_t$$

3.4. نماذج GARCH غير المستقرة IGARCH و نماذج GARCH المتكاملة الكسرية:

اقترح Engle and Bollerslev (1987) نموذجا من نوع IGARCH وهي متعلقة بحالة وجود جذر وحدوي Unit Root في سيرورة التباين الشرطي، لهذا تتميَّز بأن لها تأثير صمود في التباين Persistence Effects، وهذا يعني أن كل صدمة على التباين الشرطي

الحالي سوف تنعكس على كل القيم المستقبلية المتوقعة. إن دراسة الاستقرارية (من الرتبة الثانية) لسيرورة GARCH تقتضي بأن يكون التباين غير الشرطي مستقل بشكل تقريبي asymptotically عن الزمن، يكتب النموذج على الشكل التالي:

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}$$

مع: $\alpha_0 > 0$ ، $\alpha_i \geq 0,$ ، $i = 1,....,p,$ ، $\beta_j \geq 0,$ ، $j = 1,....,q$ و كثير الحدود

$$1 - \sum_{i=1}^{p} \alpha_i L^i - \sum_{jj=1}^{q} \beta_j L^j = 0$$

يحتوي على d ($d > 0$) جذر وحدوي و $\max(p,q) - d$ جذر خارج نطاق الوحدة. لهذه السيرورة درجة تكامل في التباين تساوي d إذا كانت $\alpha_0 = 0$ و درجة تكامل مع اتجاه عام إذا كانت $\alpha_0 > 0$. في حالة وجود نموذج IGARCH يشترط أن يكون:

$$\sum_{i=1}^{p} \alpha_i + \sum_{jj=1}^{q} \beta_j = 1$$

صدمة على التباين تنعكس على تنبؤات كل قيمها المستقبلية، دالة رد الفعل على التذبذبات تؤول إلى حد ثابت غير معدوم.

هناك نوع آخر من السيرورات غير المستقرة، يتعلق الأمـ ـر بنمـ ـاذج FIGARCH و HYGARCH الذي يعتبر مزيج بين نموذج GARCH و السيرورة الكسرية المتكاملـ ـة[1]. اقترح Baillie, Bollerslev and Mikkelsen (1996) السيرورة FIGARCH التي تُنمذِّج فقط الحالة التي يكون فيها تناقص معـ ـاملات الارتبـ ـاط علـ ـى شـ ـكل قطـ ـع زائـ ـد (hyperbolicc). في حالة GARCH(1,1) لدينا:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-i}^2 + \beta_1 h_{t-j}$$

الذي يمكن كتابته على الشكل: $(1 - \beta_1 L)h_t = \alpha_0 + \alpha_1 \varepsilon_{t-i}^2$

1- سنتطرق إلى هذا النوع من النماذج في الفصل الموالي عند دراستنا للسيرورات ذات الذاكرة الطويلة.

لدينا:

$$h_t = \frac{\alpha_0}{1-\beta_1 L} + \frac{\alpha_1}{1-\beta_1 L}\varepsilon_{t-i}^2 = \frac{\alpha_0}{\beta(1)} + \left[1 - \frac{1-\vartheta_1 L}{1-\beta_1 L}\right]\varepsilon_t^2 = \frac{\alpha_0}{\beta(1)} + \vartheta(L)\varepsilon_t^2$$

حسب السيرورة IGARCH، نجد:

$$h_t = \frac{\alpha_0}{\beta(1)} + \left[1 - \frac{1-L}{1-\beta_1 L}\right]\varepsilon_t^2$$

إذ أن:

$$\vartheta(L) = 1 - \frac{1}{\beta(L)}(1-L)$$

السيرورة FIGARCH تُدرج قوة كسرية Fractional power على عبارة الفرق الموجودة في الصيغة الأخيرة. يصبح لدينا إذن:

$$\vartheta(L) = 1 - \frac{1}{\beta(L)}(1-L)^d \ , \ 0 \le d \le 1$$

إلا أن هذه الأخيرة هي الوحيدة التي تتصف بتناقص سريع في معاملات التأخير، وهذا ما نستطيع تسميته بالذاكرة الطويلة Long Memory. بيّن Davidson (2004) أن ذاكرة هذه السيرورة تكبر كلما اقترب d من الصفر.

إذن الذاكرة هي عبارة عن دالة متزايدة مع d، وعليه يمكن النظر إلى نماذج FIGARCH كما لو أنها حالة وسيطية بين نماذج GARCH المستقرة و IGARCH، بنفس الطريقة التي نعتبر فيها أن السيرورة *I(d)* على مستوى وسيطي بين *I (0)* و *I (1)*.
لدينا:

$$\vartheta(L) = \frac{1}{\beta(L)}\left(1 + \alpha\left((1-L)^d - 1\right)\right) , \ \alpha \ge 0$$

وحسب Davidson (2004)، النماذج FIGARCH و GARCH تتعلق على التوالي بالحالات $\alpha = 0$ و $\alpha = 1$، غير أنه يمكن ملاحظة بأن المعامل d يصبح غير قابل للتعيين لما $\alpha = 0$، وهذا يؤثر سلبا على تركيب اختبارات الفروض بالنسبة لـ α، حيث نجد أن الخصائص التقاربية لمقدرات المعقولية تصبح غير محققة.

تولدت عن هذه النماذج نماذج أخرى مثل FAPARCH التي تعتبر سيرورات كسرية Fractional processes تتميز بتناقص سريع لمعاملات الارتباط الذاتي على شكل قطع زائد بحيث تسمح بوجود صفة غير متناظرة مرافقة لإشارة الأخطاء.

4.4. أنواع أخرى من نماذج ARCH:

هناك نماذج أخرى عديدة ظهرت، نذكر منها على سبيل المثال نموذج -GJR GARCH الذي اقترحه(1993) Glosten, Jagannathan and Runkle الذي يأخذ في الحسبان القدوم المفاجئ وغير المتوقع للأحداث، وهذا بإدخال متغير مفسر جديد:

$$\varepsilon_t = \eta_t h_t^{1/2}$$
$$\varepsilon_t \mid \varepsilon_{t-1} \sim N(0, h_t)$$
$$h_t = \alpha_0 + \sum_{i=1}^{p} \left(\alpha_i \varepsilon_{t-i}^2 + \gamma_i I_{\varepsilon_{t-i}<0} \varepsilon_{t-i}^2 \right) + \sum_{j=1}^{q} \beta_j h_{t-j}$$

$I_{\varepsilon_{t-i}<0}$ تمثل الدالة الثنائية (الصورية) بحيث:

$$I_{\varepsilon_{t-i}<0} = \begin{cases} 1 \text{ , if } \ \varepsilon_{t-i} < 0 \\ 0 \text{ , otherwise} \end{cases}$$

نستطيع إعطاء تعميم لنماذج GJR-GARCH بواسطة النموذج VS-GARCH المقترحة من طرف(1996-1997) Fornari and Mele حيث أن جميع العوامل تتغير حسب النظام وليس فقط عوامل مربعات الأخطاء الماضية. نقول أن السيرورة ε_t تحقق النم وذج -VS GARCH(1,1) إذا وفقط إذا كان:

$$\varepsilon_t = \eta_t h_t^{1/2}$$
$$\varepsilon_t \mid \varepsilon_{t-1} \sim N(0, h_t)$$
$$h_t = \left(\omega_{pos} + \alpha_i \varepsilon_{t-i}^2 + \beta_{pos} h_{t-1} \right)\left(1 - I_{\varepsilon_{t-i}<0}\right)$$

التباين غير الشرطي لـ . ε_t معطى بالعلاقة:

$$\sigma_\varepsilon^2 = E(\varepsilon_t^2) = \frac{(\omega_{pos} + \omega_{neg})/2}{1 - (\alpha_{pos} + \alpha_{neg})/2 - (\beta_{pos} + \beta_{neg})/2}$$

بواسطة نماذج GJR-GARCH نستطيع نمذجة نظامين للتباين الشرطي، حيث يكون اختيار نظام معين حسب إشارة الخطأ الماضي، حيث:

إذا كان $\varepsilon_{t-1} > 0$ فإن: $h_t = \alpha_0 + \alpha_{pos}\varepsilon_{t-1}^2 + \beta_1 h_{t-1}$

إذا كان $\varepsilon_{t-1} \leq 0$ فإن: $h_t = \alpha_0 + \alpha_{neg}\varepsilon_{t-1}^2 + \beta_1 h_{t-1}$

أدرج Gonzales and Rivera (1998) دالة انتقال لوجستيكية Logistic Transition function بواسطة النماذج LSTGARCH[1] ، التي – إذا وضعنا $p = q = 1$ من أجل التبسيط– تُعرِّف h_t على النحو:

$$h_t = \alpha_0 + \alpha_{pos}\varepsilon_{t-1}^2\left[\Lambda(\theta\varepsilon_{t-1})\right] + \alpha_{neg}\varepsilon_{t-1}^2\left[1 - \Lambda(\theta\varepsilon_{t-1})\right] + \beta_1 h_{t-1}$$

مع:

$$\Lambda(\theta\varepsilon_{t-1}) = \frac{1}{1 + \exp(-\theta\varepsilon_{t-1})} \ , \ \theta > 0$$

ويمكن كتابة h_t أيضا من الشكل:

$$h_t = \alpha_0 + \left(\alpha_{pos}\Lambda(\theta\varepsilon_{t-1}) + \alpha_{neg}\left[1 - \Lambda(\theta\varepsilon_{t-1})\right]\right)\varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

يشكل معامل الخطأ الماضي توفيقة خطية للعوامل المتعلقة بكل نظام، فحينما يكون ε_{t-1} كبيرا بالقيمة المطلقة وموجبا فإنه يقترب من α_{pos} وعندما يكون سالبا فإنه يكون مقاربا لـ α_{neg} .

وانطلاقا من فكرة الانتقال اللطيف بين الأنظمة المنمذجة من طرف Gonzales and Rivera (1998) بالمرور من GJR-GARCH إلى LSTGARCH، عرَّف Anderson, Nam and Vahid (1999)، انطلاقا من نماذج VSGARCH، نماذج جديدة تعرف باسم ANSTGARCH "Asymmetric Nonlinear Smooth Transition GARCH".

حيث:

$$\begin{aligned} h_t = &\left(\omega_{pos} + \alpha_{pos}\varepsilon_{t-1}^2 + \beta_{pos}h_{t-1}\right)\Lambda(\theta\varepsilon_{t-1}) \\ &+ \left(\omega_{neg} + \alpha_{neg}\varepsilon_{t-1}^2 + \beta_{neg}h_{t-1}\right)\left(1 - \Lambda(\theta\varepsilon_{t-1})\right) \end{aligned}$$

1- اختصار إلى Logistic Smooth Transition GARCH

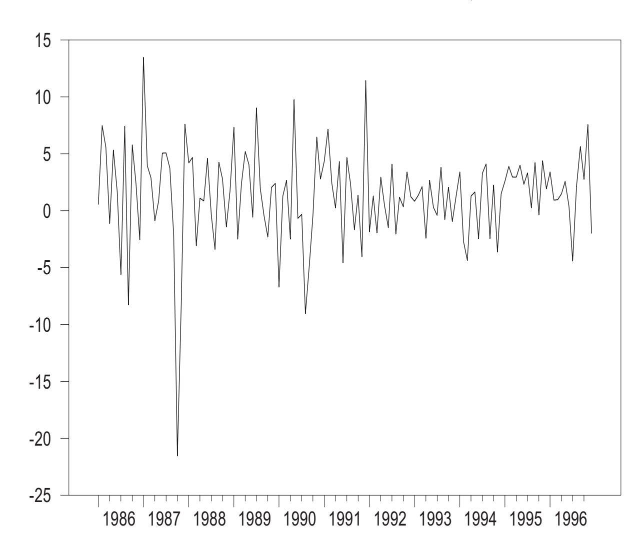
مع:
$$\Lambda(\theta\varepsilon_{t-1}) = \frac{1}{1+\exp(-\theta\varepsilon_{t-1})} \ , \ \theta > 0$$

مثال 3:

نأخذ مثالا عن مردودية مؤشر SP500 في الفترة الممتدة بين شهر يناير 1986 و ديسمبر 1996 (أنظر الشكل رقم 13). نقوم بتقدير نماذج IGARCH(1,1)، EGARCH(1,1) و GJR-GARCH(1,1) باستعمال برمجية RATS 5.04 وفق خوارزمية BFGS. نفترض أن هذه السلسلة لا تحتوي إلا على التباين الشرطي أما المتوسط (التوقع) الشرطي يتمثل في الثابتة.

```
cal 1986 1 12
all 1996:12
open data returns.xls
data(format=xls,org=cols)
*
compute gstart=2,gend=1996:12
set y = sp500
*
nonlin(parmset=meanparms) b0
frml resid = y - b0
declare series u    ;* Residuals
declare series h    ;* Variances
```

الشكل رقم (13): سلسلة مردودية مؤشر SP500



- بالنسبة للنموذج IGARCH:

نقدر النموذج وفق التعليمة التالية للحصول على النتائج المبينة أدناه:

```
NONLIN(parmset=garchparms) VC VA VB VA+VB==1.0 VA>=0.0 VB>=0.0
FRML HF = VC + VA*H{1} + VB*U{1}**2
FRML LOGL = (H(T)=HF(T)),(U(T)=RESID(T)),-.5*(log(h)+u**2/h)
LINREG(NOPRINT) Y / U
# CONSTANT
COMPUTE B0=%BETA(1)
COMPUTE VC=%SEESQ,VA=.50,VB=.50
SET H = %SEESQ
MAXIMIZE(parmset=meanparms+garchparms,METHOD=SIMPLEX,ITERS=5,NOPRINT) LOGL GSTART GEND
MAXIMIZE(parmset=meanparms+garchparms,METHOD=Bfgs,robusterrors,ITERS=100) LOGL GSTART GEND

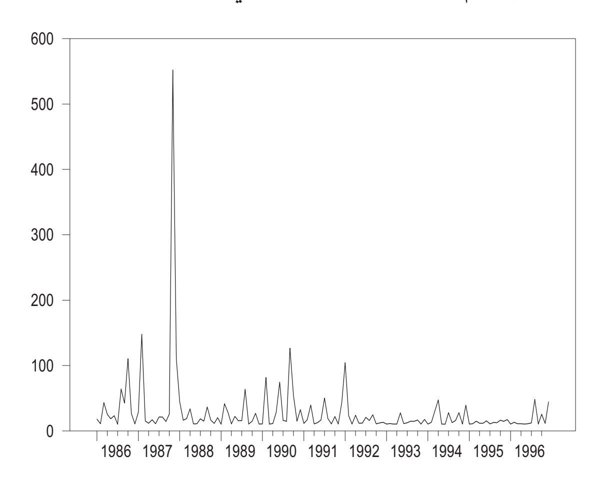
MAXIMIZE - Estimation by BFGS
Convergence in    11 Iterations. Final criterion was  0.0000003 <  0.0000100
Monthly Data From 1986:02 To 1996:12
Usable Observations    131
Function Value                   -254.37593831

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  B0                        1.737398       0.305473      5.68757  0.00000001
2.  VC                       10.349904       2.113347      4.89740  0.00000097
3.  VA                       -0.000000       0.000000      0.00000  0.00000000
4.  VB                        1.000000       0.000000      0.00000  0.00000000
```

نمثل بيانيا دالة التباين الشرطي:

```
graph 1
#H
```

الشكل رقم (14): التباين الشرطي IGARCH(1,1)



- بالنسبة للنموذج EGARCH:

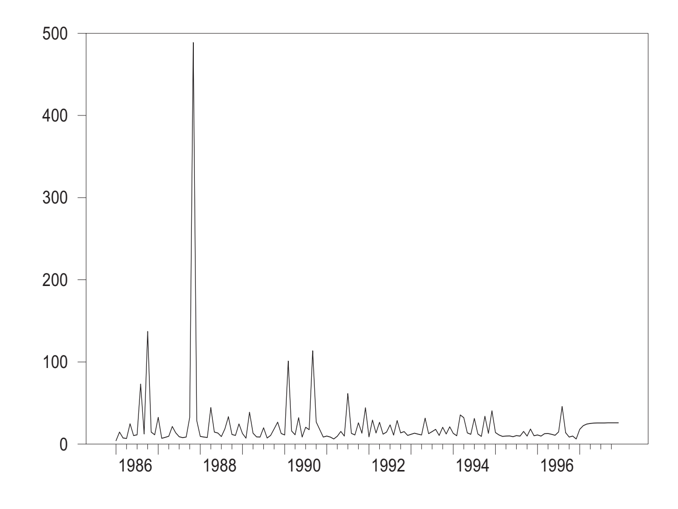
```
NONLIN(PARMSET=GARCHPARMS) VC VA VB VD
FRML G = ABS(U(T)/SQRT(H(T))) - SQRT(2.0/%PI) - VD*U(T)/SQRT(H(T))
FRML HF = EXP(VC + VA*LOG(H{1}) + VB*G{1})
FRML LOGL = (H(T)=HF(T)),(U(T)=RESID(T)),-.5*(LOG(h)+u**2/h)
LINREG(NOPRINT) Y / U
# CONSTANT
COMPUTE B0=%BETA(1)
COMPUTE VC=LOG(%SEESQ),VA=.05,VB=.05,VD=.05
SET H = %SEESQ
MAXIMIZE(parmset=meanparms+garchparms,METHOD=SIMPLEX,START=INIT,ITERS=5,NOPRINT) LOGL GSTART GEND
MAXIMIZE(parmset=meanparms+garchparms,METHOD=Bfgs,robusterrors,START=INIT,ITERS=100) LOGL GSTART GEND

MAXIMIZE - Estimation by BFGS
Convergence in    29 Iterations. Final criterion was  0.0000021 <  0.0000100
Monthly Data From 1986:02 To 1996:12
Usable Observations    131
Function Value                   -244.77402298

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  B0                      1.1804906467 0.3279196724      3.59994  0.00031829
2.  VC                      2.4116589911 0.5281836946      4.56595  0.00000497
3.  VA                      0.1285267503 0.2037696011      0.63075  0.52820697
4.  VB                      0.2917150437 0.2240524457      1.30199  0.19291828
5.  VD                      2.0690799052 1.2984821692      1.59346  0.11105695
```

نمثل بيانيا دالة التباين الشرطي:

الشكل رقم (15): التباين الشرطي EGARCH(1,1)



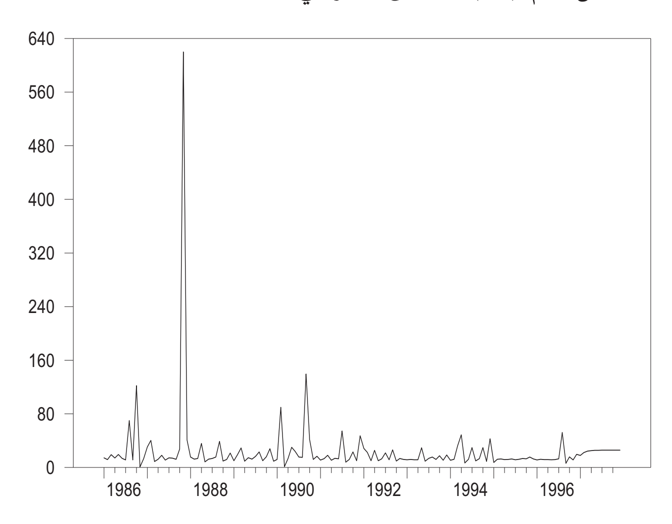
- بالنسبة للنموذج GJR-GARCH:

```
NONLIN(PARMSET=GARCHPARMS) VC VA VB VD
FRML HF = VC + VA*H{1} + VB*U{1}**2 + %IF(U{1}<0.0,VD*U{1}**2,0.0)
FRML LOGL = (H(T)=HF(T)),(U(T)=RESID(T)),-.5*(LOG(h)+u**2/h)
LINREG(NOPRINT) Y / U
# CONSTANT
COMPUTE BO=%BETA(1)
COMPUTE VC=%SEESQ,VA=.05,VB=.05,VD=.05
SET H = %SEESQ
MAXIMIZE(parmset=meanparms+garchparms,METHOD=SIMPLEX,START=INIT,ITERS=5,NOPRINT) LOGL GSTART GEND
MAXIMIZE(parmset=meanparms+garchparms,METHOD=Bfgs,robusterrors,START=INIT,ITERS=100) LOGL GSTART GEND

MAXIMIZE - Estimation by BFGS
Convergence in    22 Iterations. Final criterion was  0.0000000 <  0.0000100
Monthly Data From 1986:02 To 1996:12
Usable Observations    131
Function Value                    -247.30359439

   Variable                    Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  BO                      1.55125492   0.66332438      2.33861  0.01935579
2.  VC                     13.06310540   2.69443921      4.84817  0.00000125
3.  VA                     -0.13180774   0.12710628     -1.03699  0.29974125
4.  VB                      0.22127433   0.20595603      1.07438  0.28265395
5.  VD                      0.92419573   1.13644414      0.81323  0.41608354
```

الشكل رقم (16): التباين الشرطي GJR-GARCH(1,1)



الفصل التاسع

طرق غير خطية

في تحليل السلاسل الزمنية

الفصل التاسع

طرق غير خطية في تحليل السلاسل الزمنية

تناولنا في الفصل السادس شكلا يتمثل في نماذج السلاسل الزمنية الخطية، التي تعتمد في تفسيرها للظاهرة في اللحظة الحالية على المتوسطات المرجحة للملاحظات الماضية والأخطاء العشوائية، لكن هذا النوع من النماذج لا يأخذ بعين الاعتبار الحركية غير المتناظرة و غير الخطية في السلاسل. سنتطرق في هذا الفصل إلى طرق أخرى في تحليل السلسلة الزمنية وتتمثل في الديناميكية غير الخطية للسلسلة، حيث نعالج حالة وجود بنية الشواش و نعطي مفهوما لها ثم نتطرق إلى النماذج ذات الذاكرة الطويلة و في الأخير نعطي نظرة حول السيرورات غير المعلمية وطريقة النواة.

1. الشواش Chaos: التفسير التحديدي (أو الثابت) للتقلبات

1.1. مفهوم النظام المشوش:

تعتبر التقلبات الدورية لأي سلسلة زمنية داخلية (يفسر بالنظام ذاته)، فالحركية (الديناميكية) الجوية لم تكن أبدا عشوائية بل تحديدية (ثابتة) وأن صعوبة التنبؤ ناجم عن سرعة التأثر بالشروط الابتدائية Initial conditions[1]. هناك تشابه بين السلاسل المالية و الجوية بسبب الصعوبات التي يمكن أن تعترض الإحصائي عند التنبؤ بمؤشر البورصة، فهي ناتجة عن وجود مسارات معقدة وغير قابلة للتنبؤ وهذا ما لا نلاحظه إلا في النظام المشوش.

للتعريف بنظرية الشواش، نستعمل المعادلة التالية:

$$Y_t = f(Y_{t-1}),\ \ Y_0 \in D \subset R^k$$

حيث Y_0 يعبر عن الشرط الابتدائي.

1- الشروط الابتدائية هي تلك القيم الابتدائية الخاصة بالنظام الديناميكي (الحركي) والتي تنتمي إلى المجال [0,1].

حسب نظرية الشواش، تحدث الصدفة (المسارات المعقدة) بسبب سرعة التأثر بالشروط الابتدائية. كذلك قد يؤدي اضطراب صغير في الشرط الابتدائي إلى نتائج مختلفة تماما حيث يمكن تفسير هذه الظاهرة وذلك باعتبار أن خطأ صغير يسبب في خطأ كبير. في نموذج مشوش Chaotic model، الحركة العشوائية ليست إلا نتيجة وبالتالي طبيعة النظام المسبب في هذه الحركة تحديدي بحث.

النظام الحركي المشوش هو نظام حركي تحديدي Deterministic Dynamic System سريع التأثر بالشروط الابتدائية. في هذا النوع من النظام، لا يمكن التنبؤ بالظاهرة على فترات طويلة.

نقول أن نظام معرف بالتطبيق F أنه مشوش إذا كان:

- دالة الارتباط الذاتي تنعدم من أجل فجوة زمنية (تباطؤ) قصيرة،
- جاذب النظام Attractor في فضاء الأطوار Phase Space[1] يعتبر جاذبا غريبا،
- سريع التأثر بالشروط الابتدائية.

ملاحظات:

- سيرورة مشوشة هي تقريبا مرتبة ordered بمعنى أنها تخضع إلى معادلة تحديدية في فضاء الأطوار.
- السيرورة تحديدية بمعنى أنها إذا كانت الشروط الابتدائية تتكرر بصفة منتظمة، فإن تطور النظام عبر الزمن يبقى دائما نفسه.
- سيرورة مشوشة تعتبر غير دورية.

1- يعتبر الجاذب مجموعة جزئية من نقاط حولها تقترب كل مسارات النظام الحركي و فضاء الأطوار هو الفضاء الذي من خلاله تمثل المحاور إحداثيات المتغيرات الحركية المستقلة عن النظام.

2.1. اختبارات الكشف عن ظاهرة مشوشة:

1.2.1. اختبار بعد الارتباط Correlation Dimension Test:

اقترح (1983) Grassberger and Procaccia اختبارا يعتمد على بعد الارتباط الذي يعتبر بديلا لاختبار Hausdorff.

لتكن $\{Y_t\}$ سلسلة زمنية ولنعتبر مجموعة النقاط $\{Y_t, t=1,2,......,T\}$ على الجاذب Attractor. نفترض أن جاذب النظام مشوش Chaotic ونأخذ نقطتين $\{Y_i, Y_j\}$ لهما نفس المسار بعيدين في الزمن. نظرا لتبعيتهما للشروط الابتدائية، هاتان النقطتان غير مرتبطتين حركيا لأن خطأ صغير متعلق بتحديد النقطة الابتدائية قد يؤدي إلى وجود موقع مختلف للنقطة الثانية، فالنتيجة هذه تعتبر منطقية باعتبار أن التبعية للشرط الابتدائي تؤدي إلى وجود مسارات متباعدة، فتقع هذه النقاط في فضاء محدود بحيث قد تكونان مرتبطتين في الجاذب.

نقيس هذا الارتباط الفضائي بتكامل الارتباط Correlation Integral:

$$C(\varepsilon, m, T) = \frac{1}{T_m(T_m - 1)} \sum_{i,j=1}^{T} H\left(\varepsilon - \left\|Y_i - Y_j\right\|\right), \quad i \neq j$$

حيث $T_m = T - m + 1$ ، $\left\|Y_i\right\| = \max\left(\left|Y_i\right|\right)$ المعيار الاقليدي و H هي دالة Heaviside:

$$H\left(\varepsilon - \left\|Y_i - Y_j\right\|\right) = \begin{cases} 1, & si: \left\|Y_i - Y_j\right\| \leq \varepsilon \\ 0, & \left\|Y_i - Y_j\right\| > \varepsilon \end{cases}$$

كما هو محقق بفضل معادلة تكامل الارتباط، النقطتان Y_i و Y_j مرتبطتان فضائيا إذا كانت المسافة الاقليدية أقل من ε. هذه الأخيرة هي المسافة القصوى بين زوجين من النقاط. من أجل القيم المتزايدة لـ ε ، $C(\varepsilon, m, T)$ يكبر ويصبح مساويا إلى 1 انطلاقا من بعض قيم ε كبيرة بشكل كاف ($0 \leq C(\varepsilon, m, T) \leq 1$).

برهن كل من Denker and Keller (1986) و Brock and Dechert (1988a) أنه من أجل أغلب الشروط الابتدائية، $T \to \infty, C(\varepsilon,m,T) \xrightarrow{d} C(\varepsilon,m)$ (أي $C(\varepsilon,m,T)$ تؤول بتوزيع احتمالي إلى $C(\varepsilon,m)$). لدينا[1]:

$$C(\varepsilon,m) = \left[C(\varepsilon,1)\right]^k$$

يقيس تكامل الارتباط جزءا من أزواج الشعاعين $(y_i, y_{i+1},...., y_{i+m+1})$ و $(y_j, y_{j+1},...., y_{j+m+1})$ بحيث أن البعد بينهما أقل من ε.

- إذا كان ε مختارا بحيث كل أزواج الأشعة تحقق الشرط $\left\|(y_i, y_{i+1},...., y_{i+m+1}),(y_j, y_{j+1},...., y_{j+m+1})\right\| = \max\limits_{m\in[0,k-1]}\left\|y_{i+m} - y_{j+m}\right\| < \varepsilon$ ، فإن $C(\varepsilon,m,T) = 1$ (وجود ارتباط فضائي)
- إذا كان ε مختارا بحيث أن الشرط غير محقق تماما، فهذا يعني أن $C(\varepsilon,m,T) = 0$. إذن ليس هناك ارتباط.

يعرف بعد الارتباط رياضيا كما يلي:

$$D_c = \lim_{c\to 0}\frac{\ln C(\varepsilon)}{\ln \varepsilon}$$

تطبيقيا، إذا كانت لدينا سلسلة زمنية Y_t حيث $t = 1,....,T$ ، فمن الممكن تقدير هذا البعد ببناء m-تاريخ لنظام في فترات متقطعة:

$$Y_i^m = (Y_i, Y_{i+1},...., Y_{i+m+1})$$

حيث m يسمى بـ . Embedding dimension.

يتم حساب تكامل الارتباط $C(\varepsilon,m)$ على m-تاريخ للسلسلة ويرتبط بعدد عناصر هذا الشعاع m. من أجل كل القيم الصغيرة لـ ε ، برهن Grassberger and Procaccia (1983) أن $C(\varepsilon,m)$ يتزايد بصفة أسية:

$d_c = \lim\limits_{c\to 0}\dfrac{\ln C(\varepsilon,m)}{\ln \varepsilon}$، إذن:

$C(\varepsilon,m) = \varepsilon^{d_m} \Leftrightarrow \ln C(\varepsilon,m) \approx \ln \varepsilon^{d_k} \Leftrightarrow \ln C(\varepsilon,m) \approx dm \ln \varepsilon$.

1- Brock and Break (1991)

يتزايد $C(\varepsilon,m)$ بوتيرة ε^{d_k}. عندما يكبر البعد m "Embedding Dimension"، نحصل، من أجل كل قيمة لـ m، على البعد d_m. خلال هذه العملية، d_m تستقر عند القيمة $\hat{d}_c$ والتي تعبر عن بعد ارتباط الجاذب "Correlation dimension of attractor":

$$\hat{d}_c = \lim_{k\to\infty} d_m$$

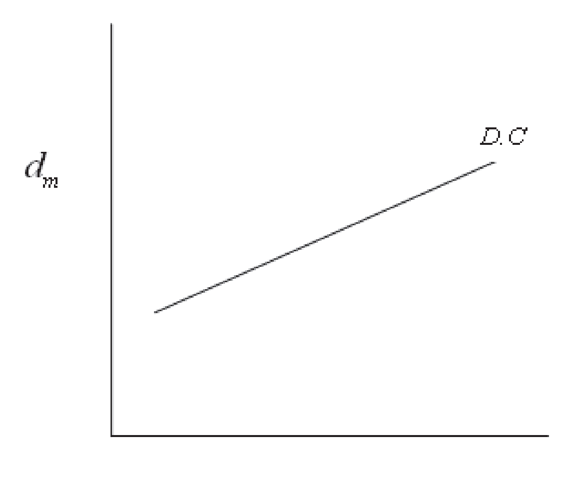
أو:

$$\hat{d}_c = \lim_{m\to\infty}\left[\frac{d\ln C(\varepsilon,m)}{d\ln\varepsilon}\right]$$

إن طريقة بعد الارتباط تشكل أداة فعالة يمكن استخدامها من أجل التمييز بين سيرورة تحديدية و سيرورة عشوائية.

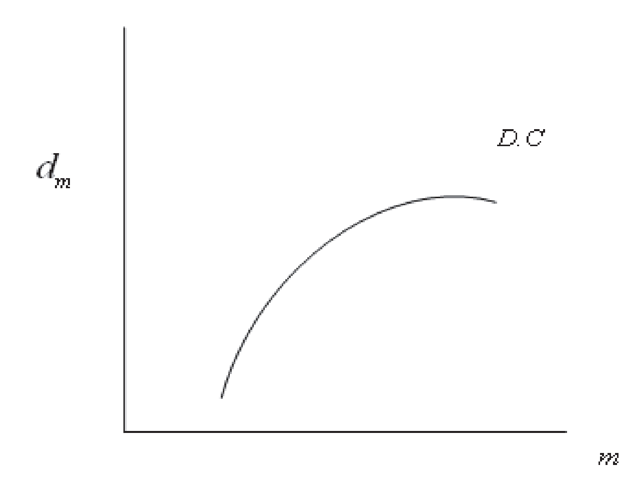
1. في حالة ما إذا كانت السلسلة تتضمن متغيرات عشوائية بحتة، فإن هذه الأخيرة لها بعد غير منتهي ونتيجة لذلك، يتم ملأ فضاء الأطوار العشوائية التي لها بعد m. سنحصل على العلاقة $\lim_{m\to\infty} d_m = d_c = \infty$. في حالة سيرورة عشوائية، لا يمكن أن نحصل على الإشباع Saturation، لأن d_m لا تستقر عندما تكبر m و يجب أن يتزايد بعد الارتباط على وتيرة واحدة مع البعد "Embedding dimension" m:

الشكل رقم (1): $\hat{d}_c$ لسيرورة عشوائية



2. في حالة ما إذا كانت المعطيات ممثلة بنظام تحديدي deterministic، d_m تستقر عند مستوى $\hat{d}_c$ عندما تكبر m ونحصل على بعد معين. بمعنى آخر، d_m تصبح مستقلة عن m من أجل $m \geq 2T+1$ حيث T يمثل بعد النظام الذي تخضع له المعطيات:

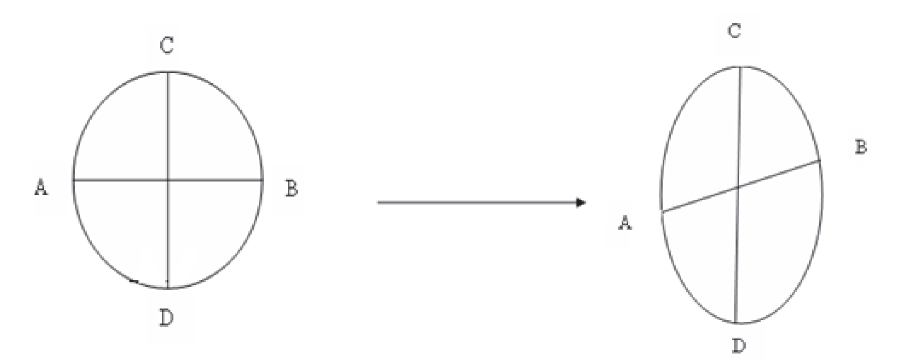
الشكل رقم (2): $\hat{d}_c$ لسيرورة تحديدية



1.2.1. اختبار أس Lyapunov "Lyapunov Exponent Test":

تقيس أسس Lyapunov تباعد أو تقارب المسارات انطلاقا من نقاط قريبة جدا، فوجود أسس موجبة (تباعد المسارات) يشير إلى فقدان القدرة التنبؤية خلال فترات زمنية معينة. إذا كان الانحراف الابتدائي بين نقطتين قريبتين جدا يتزايد أسيا،، فحتى ولو كان الشرط الابتدائي مقاسا بدقة تامة فالتنبؤ المستقبلي للنظام على المدى القصير غير ممكن. بما أن النقطة الابتدائية معروفة بدقة تامة، فيمكن اعتبار أن القيمة الحقيقية الابتدائية تقع في مجال بعده n تتوسطه هذه النقطة، ففي هذه الحالة، سوف يتبدل شكل محاور هذا المجال والتي عددها n خلال فترة زمنية معينة و سوف يفقد خصائص التعامد (الشكل رقم (3)). يتحول المجال الابتدائي Initial sphere إلى مجسم ناقص Ellipsoid ذي بعد n.

الشكل رقم (3): " *stretching* "و " *contracting* " في نظام حركي



إن أثر "stretching" يعني أن نقطتين ابتدائيتين قريبتين من بعضهما البعض في الدورة الأصلية Original Cycle سوف يتباعدان أسيا على الجاذب أي أن أسا موجبا لـ . Lyapunov يقيس نسبة تباعد نقطتين ابتدائيتين، أما أثر " contracting " فهو يفسر على أن النقطتين تصف مسارين يتقاربان بعد فترة زمنية معينة.

ونتيجة لذلك، إذا كانت المسافة بين هاتين النقطتين تتزايد بشكل أسي (سرعة التأثر بالشروط الابتدائية)، فإن هناك صفة شواش chaotic character. كذلك، بوضع λ_1 أكبر أس Lyapunov، لدينا المعيار التالي:

استقرار إذا كان $\lambda_1 < 0$

شواش إذا كان $\lambda_1 > 0$

بصفة عامة، يمكن حساب كل أسس Lyapunov باستعمال المعادلة التالية:

$$\lambda_i = \lim_{T\to\infty} \frac{1}{T} \log\left\{2\Gamma_i^{(T)}\right\}$$

حيث $\Gamma_i^{(T)}$ تمثل القيم الذاتية للمصفوفة الجاكوبية (Jacobian Matrix) للنظام.

يلاحظ هنا أن أس Lyapunov معرف كنهاية تقاربية Asymptotic limit ($T \to \infty$) حتى نتمكن من وصف السلوك طويل المدى للنظام. لكي يكون الجاذب مشوشا، يجب أن يكون هناك على الأقل أس واحد لـ . Lyapunov موجب.

يمكن حساب أسس Lyapunov وذلك باستخدام خوارزمية Wolf and al, Wolf (1985) وأيضا خوازميات أخرى كخوارزميات Schreiber (1993) أو Nycka and al (1992) التي أعطت نتائج قوية وفعالة حتى في ظل وجود التشويش The noise، فحساب أسس Lyapunov يمكن أيضا أن يعطينا صيغة رياضية أخرى لبعد Hausdorff. يعرف بعد Lyapunov والذي نرمز له بـ . D_L كما يلي:

$$d_L = m + \frac{\sum_{i=1}^{m} \lambda_i}{|\lambda_{m+1}|}$$

بشرط أن يكون $\sum_{i=1}^{m} \lambda_i > 0$ و $\sum_{i=1}^{m+1} \lambda_i < 0$.

في حالة جاذب مشوش ذي ثلاث أبعاد، لدينا العلاقة التالية:

$$d_L = 2 + \frac{\lambda_1}{|\lambda_3|}$$

بصفة عامة، بعد Lyapunov أكبر من أو يساوي بعد Hausdorff.

في الجانب التطبيقي، نعطي مسافة قصوى لا يمكن اجتيازها وفي حالة اجتيازها، المسار المتبع سابقا لا بد أن يستبدل بمسار جديد وهكذا... وإذا كان جاذب النظام مشوشا، فالحساب يجب أن يؤول إلى قيمة مستقرة λ_1.

يحتاج تطبيق هذه التقنية إلى اختيار عدة معالم:

- البعد "Embedding dimension".
- الوقت الذي يستغرق من أجل تتبع مسار معين قبل أن يستبدل بمسار جديد.
- المسافة القصوى بين النقطة المدرجة على المسار المرجعي ونقطة المسار الذي سنتبعه خلال الفترة المحددة سابقا.
- المسافة القصوى بين هاتين النقطتين.

إذا كانت الخوارزمية من الجانب النظري لا تتأثر باختيار هذه المعالم، ففي الجانب التطبيقي أيضا ليس هناك تأثير. من المهم إذن القيام بعدة محاولات والاحتفاظ بالقيم التي

تقود إلى نتائج مستقرة. في حين أن Wof and al (1985) اقترحا، بواسطة محاكاة، عدة قواعد تسمح بمساعدتنا على اختيار المسافات القصوى والدنيا المقبولة:

- المسافة القصوى بين نقطتين لا يجب أن تتعدى 10% من مجال السلسلة.
- المسافة الدنيا بين نقطتين مختارة بصفة عامة على أنها تشكل 10% من المسافة القصوى.

أما فيما يتعلق باختيار الزمن الذي نحتاجه لتتبع مسار قبل استبداله بمسار جديد، ينبغي القيام بعدة محاولات بغية إيجاد قيم مستقرة نسبيا لأس Lyapunov.

مثال 1:

في هذا المثال، سنطبق هذين الاختبارين على سلسلة مردودية SP500 خلال الفترة الممتدة بين 1968/01/02 و 1996/06/12. نقوم أولا بتقدير بعد الارتباط Correlation dimension والنتائج تظهر في الجدول التالي:

الجدول (1): نتائج تقدير بعد الارتباط

15	10	5	2	*m*
3.21	2.31	1.25	0.52	C.D

نلاحظ من خلال الجدول (1) أن بعد الارتباط يتزايد مع البعد " Embedding dimension" ولكن أقل سرعة من البعد "Embedding dimension" وهذا يعني أن هذه المقدرات لا تقترب نحو قيمة مستقرة (أي وجود بنية مشوشة).

تجدر الإشارة هنا إلى أنه بالرغم من أن للنظام بعد محدود، فلا يمكن القول أن الأمر يتعلق بسيرورة مشوشة، فينبغي تطبيق تقنية أس Lyapunov بالرغم من أن طريقة بعد الارتباط تعطي إشارة إلى وجود طبيعة تحديدية أو عشوائية للسيرورة غير أنها لا تشكل لا شرطا ضروريا ولا كافيا لكي يكون النظام مشوشا.

يعطي الجدول (2) نتائج اختبار أس Lyapunov. يتغير البعد " Embedding dimension" من 3 إلى 10 خلال الفترتين 10 أيام و 40 يوما:

الجدول (2): حساب أكبر أس Lyapunov

الفترة الزمنية	$m=3$	$m=4$	$m=5$	$m=10$
10	0.0174	0.0241	0.0242	0.0105
40	0.0047	0.0065	0.0072	0.0022

نلاحظ أولا أن كل أسس Lyapunov موجبة توحي بوجود جاذب مشوش، إضافة إلى ذلك، يعتبر أكبر أس Lyapunov موجبا وهذا إشارة إلى أن القدرة التنبؤية لسلسلة المردودية ضعيفة، كلما كانت قيمة أس Lyapunov كبيرة كلما كانت القدرة على التنبؤ ضعيفة.

نلاحظ من جهة أخرى أن قيمة أس Lyapunov مستقرة نسبيا وخاصة عند الأبعاد "Embedding dimension" 3، 4 و5. زيادة على ذلك، تتناقص قيمة الأس عندما تكبر الفترة الزمنية التي نستغرقها لتتبع مسار معين قبل أن يستبدل بمسار آخر.

يمكن القول إذن أن هذين الاختبارين يشيران إلى أن طبيعة السلسلة تحديدية، أما قيم أسس Lyapunov الضعيفة المتحصل عليها تجعلنا نتساءل على المعنوية الفعلية لهذه المقدرات: هل تختلف معنويا عن الصفر؟ يبقى هذا السؤال دون جواب لأن حساب أسس Lyapunov لا يعبر عن اختبار إحصائي بمعنى الكلمة ونتيجة لهذه الانتقادات، اقترح اختباران آخران: اختبار البواقي Residual test و اختبار "المزج العشوائي" Mixture test. تم تطبيق أول اختبار مقترح من طرف Brock (1986) على سلسلة بواقي تقدير نموذج ARMA. نذكر فقط أنه إذا كانت السيرورة تحديدية، فيجب أن نحصل على نفس النتائج مثل السلسلة الأصلية، أما فيما يخص اختبار المزج العشوائي، إذا كانت السيرورة قيد الدراسة مشوشة، فإن البعد المقدر على السلسلة الممزوجة يجب أن يكون مرتفعا و

أس Lyapunov لا بد أن يكون ضعيفا أو سالبا. بصفة عامة، وفق (1998) .Mignon، تقود نتائج هذين الاختبارين إلى رفض فرضية الشواش التحديدي.

3.1. مشكل التشويش وصعوبة تحديد طبيعة السيرورة في الأسواق المالية:

تتميز دالة الارتباط الذاتي و الدالة الطيفية للأنظمة المشوشة و العشوائية بنفس الخصائص. بحكم الحساسية للشروط الابتدائية، تؤول معاملات الارتباط الذاتي الزمنية للسيرورة المشوشة إلى الصفر على المدى الطويل و النظام غير قابل للتنبؤ. في نموذج عشوائي، تباعد المسارات يحتاج إلى ظهور صدمة خارجية[1]، أما في حالة النموذج المشوش، يحدث هذا التباعد عند ظهور صدمة داخلية[2]. في هذه الحالة، التباعد ناتج عن سرعة التأثر بالشروط الابتدائية و المسارات تتباعد دائما على الجاذب بمعدل أسي.

قد يكون للظاهرة سلوك مختلف تماما في حالة وجود تشويش Noise، تشويش قوي (تباين مرتفع) يمنع معرفة ما إذا كانت الحركية المناسبة ساكنة أم لا. في الأسواق المالية، يجعل التشويش المشاهدات غير جيدة مما يسمح بمشاهدة المبادلات في السوق لأنه إذا لم يكن هناك تشويش في الأسواق، فيلجأ الوكلاء The agents إلى تنويع محفظتهم بشكل عقلاني لتقدير مخطر المردودية المتوقعة، وبالتالي تكون هناك صفقات.

لتكن سلسلة زمنية $Y_t = f(Y_{t-k}, Y_{t-k+1}, ..., Y_{t-1})$ حيث $t = 1,2,...,T$ و k هي البعد "Embedding Dimension".

- إذا كانت $Y_t \ \ t \geq 1$, السيرورة التي نبحث على تحديدها و W_t السلسلة الحقيقية في حالة وجود تشويش تجميعي Additive Noise، إذن:

$$W_t = Y_t + \varepsilon_t$$

- في حالة تشويش ديناميكي أو حركي (جدائي)، لدينا:

$$W_t = f(W_{t-k}, ..., W_{t-1}) + \varepsilon_t$$

1- ظهور معلومات غير متوقعة من طرف الوكلاء.

2- تقلبات تحدث من النظام نفسه.

في الميدان التطبيقي، عندما ندرس سلاسل مؤشرات البورصة، تكون نماذج التشويش جد مصطنعة، مثل التوفيق بين التشويش التجميعي والحركي، فكل الطرق المقترحة إلى حد الآن لتخفيض التشويش تطبق فقط على جزء التشويش التجميعي وليس على التشويش الديناميكي. لهذا السبب، يعتبر تحديد طبيعة السيرورات التي تخضع لها مردوديات البورصة صعبا للغاية.

السيرورات الأكثر استعمالا لنمذجة السلاسل المالية تنقسم إلى نوعين:

- سيرورة عشوائية: $Y_t = \underbrace{\phi_0 + \phi_1 Y_{t-1}}_{\text{جزء تحديدي (ذو أصل عشوائي)}} + \underbrace{\varepsilon_t}_{\text{جزء عشوائي}}$ حيث ε_t تشويش أبيضWhite Noise.

- سيرورة مشوشة عشوائية: $Y_t = \underbrace{f(Y_{t-1})}_{\text{جزء تحديدي (نظام تحديدي معقد مثل نظام Hénon)}} + \underbrace{\varepsilon_t}_{\text{جزء عشوائي}}$

إذا كان تحديد السيرورة خاطئا، فالتنبؤات المتحصل عليها لن تكون جيدة، فإذا كانت سلسلة مؤشرات البورصة تخضع لسيرورة مشوشة و عشوائية في آن واحد، فإنه يمكن الجمع بين معادلة التباطؤ المشوشة التحديدية لـ Mackey-Glass[1] . و التشويش (الجزء العشوائي) الذي قد يكون تباينه الشرطي غير متجانس. النموذج يكتب كما يلي[2]:

$$Y_t = \alpha \frac{Y_{t-\tau}}{1+Y_{t-\tau}^2} - \delta Y_{t-1} + \varepsilon_t$$

حيث τ يمثل معلم التباطؤ و ε_t هو الخطأ العشوائي.

1- معادلة Mackey-Glass (1977) هي نظام بعدي غير محدود ولكن بعد التجاذب يتغير عندما يتغير معلم التباطؤ τ . بين Kyrstou and Terraza (2002) أنه من أجل $\tau = 1$ ، البعد أكبر من أو يساوي 7.

2- Kyrstou and Terraza (2002)

إذا كانت السلسلة المالية تتضمن تغيرات موسمية، فإنه يمكن كتابة نموذج Mackey-Glass الموسمي كما يلي[1]:

$$Y_t = \alpha \frac{Y_{t-\tau}}{1+Y_{t-\tau}^{\tau}}\left(1+\sin\omega(t-\tau)\right)-\delta Y_{t-1}+\varepsilon_t$$

حيث تشير ω إلى فترة التذبذب الموسمي $\frac{2\pi}{\omega}$ في السيرورة. تعرف التذبذبات الموسمية كما يلي $\omega=\frac{2\pi}{s}$ مع s هي الدورة ($s=1$ في حالة السلسلة السنوية، $s=2$ إذا كانت السلسلة سداسية، $s=4$ إذا كانت فصلية، $s=12$ إذا كانت شهرية،...الخ).

مثال 2:

نختبر وجود الشواش التحديدي في السلسلة اليومية لمردودية مؤشر CAC40 خلال الفترة الممتدة بين 1987/09/07 و 1999/05/28 ثم تقدير النموذج الملائم لهذه السلسلة.

الجدول (3): نتائج اختبار BDS على سلسلة المردودية

σ/ε		m
1	0.5	
5.6170	4.2522	2
7.3283	5.5562	3
8.9349	6.9679	4
10.235	8.7084	5

من خلال الجدول (3)، نلاحظ أن سلسلة مردودية مؤشر CAC40 تتميز بارتباط حيث نرفض فرضية *i.i.d* باعتبار أن إحصائيات BDS أكبر تماما من القيمة المجدولة للتوزيع الطبيعي 1.96.

يعطي الجدول (4) نتائج تقدير بعد الارتباط Correlation Dimension لسلسلة المردودية:

1- Kyrstou and Terraza (2008)

الجدول (3): نتائج تقدير بعد الارتباط

10	9	8	7	6	5	4	3	2	m
6.522	6.244	5.878	5.477	5.091	4.562	3.794	2.928	1.918	C.D

يتضح جليا أن بعد الارتباط يتزايد مع البعد "Embedding Dimension" ولا تقترب من قيمة مستقرة وهذا معروف بالنسبة لخصائص السيرورة العشوائية. نلاحظ أيضا أن مقدر بعد الارتباط مرتفع جدا، فمن الصعب إذن التمييز بين سيرورة عشوائية بحتة و سيرورة عشوائية مشوشة.

لتحديد السلوك الحركي لسلسلة CAC40، نقوم بحساب أس Lyapunov على المردودية بتطبيق خوارزمية Gençay and Dechert (1992) التي ترتكز على ما يسمى بـ . "Feedforward Neural Networks".

الجدول (4): تقدير أسس Lyapunov لمردودية CAC40.

SIC	$\times 10^{-7}$) MSE (	$\lambda(2)$	$\lambda(1)$	m
-17.0686	0.38150	-0.96070	-0.0183	1
-17.3208	0.29336	-1.05136	0.01341	2
-17.1031	0.36092	-1.10709	0.04474	3
-17.3284	0.28511	-1.04743	0.05352	4
-17.3266	0.28264	-1.04877	-0.01775	5
-17.3155	0.28279	-1.04815	-0.01860	6
-17.2871	0.28791	-1.02332	0.05006	7
-17.3030	0.28042	-1.05714	0.01201	8
-17.2862	0.28217	-1.06187	-0.01789	9
-17.3145	0.27146	-0.99168	-0.01933	10
-17.3076	0.27047	-0.95933	0.01551	11
-17.2627	0.27993	-1.06067	0.00218	12
-17.2423	0.28272	-1.04814	-0.00096	13
-17.2388	0.28076	-1.05972	0.00197	14
-17.2584	0.27242	-0.95444	0.00182	15
-17.2625	0.26850	-0.97029	0.00295	16
-17.2122	0.27938	-1.05679	0.00215	17
-17.1994	0.28004	-1.05949	0.00227	18
-17.2267	0.26964	-0.96569	0.00313	19
-17.1847	0.27827	-1.04317	-0.00155	20

في العمود الثاني والثالث، لدينا أكبر أسس Lyapunov $\lambda(1)$ و $\lambda(2)$، أما العمودان الأخيران يعطيان متوسط مربع الخطأ (MSE) و معيار Schwarz (SIC) على الترتيب. أحسن أس Lyapunov هو الذي يصغر العيارين MSE و SIC. من خلال الجدول (4)، نلاحظ أن معيار MSE يأخذ القيمة الصغرى عندما تكون $m = 6$. في هذه الحالة $\lambda(1) = 0.002954$ و $\lambda(2) = -0.97029$. من جهة أخرى، نحصل على القيمة الصغرى لـ SIC عندما تكون $m = 4$. أسس Lyapunov في هذه الحالة هي $\lambda(1) = 0.005352$ و $\lambda(2) = -1.04743$. في كلتا الحالتين، $\lambda(1)$ موجب و $\lambda(2)$ سالب. بناء على ذلك، لا يتضح جليا ما إذا كانت هذه السيرورة عشوائية بحتة أم لا. التفسير الممكن هو أن السلاسل المالية يمكن أن تتضمن بنية مشوشة ذات تباين شرطي غير متجانس. لاختبار هذه الفرضية، نقوم بتقدير نموذج Mackey-Glass (MG) المشوش التحديدي مع خطأ GARCH(1,1) على سلسلة مردودية مؤشر CAC40.

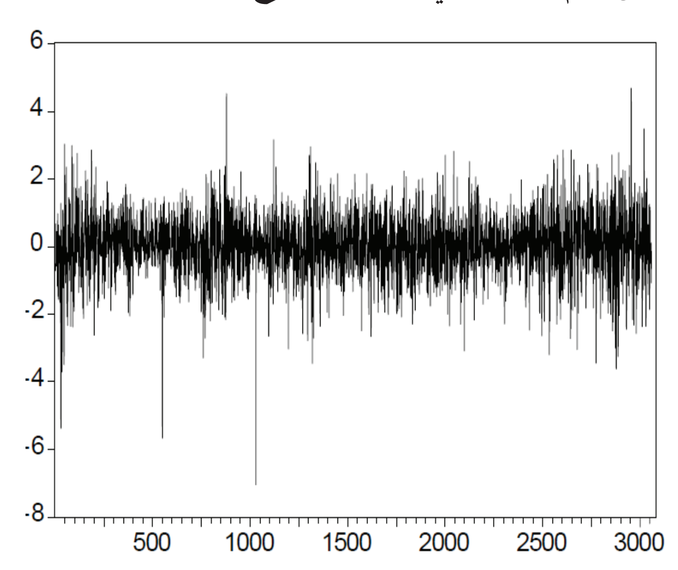
بعد تقدير نموذج Mackey-Glass ($\tau = 1$) باستخدام Eviews 5.0، نلاحظ أن بواقي التقدير ذات تباين شرطي غير متجانس لأن إحصائية ARCH-LM التي تساوي 39.44 أكبر تماما من القيمة المجدولة لتوزيع χ^2 بدرجة حرية 1 ونسبة معنوية 0.05. إذن نقدر النموذج MG-GARCH(1,1). النتائج مبينة في الجدول (5) التالي:

الجدول (5): تقدير نموذج MG-GARCH(1,1)

المعالم	القيم المقدرة
$\hat{\alpha}$	187.46 (145.10)
$\hat{\delta}$	187.38 (145.20)
$\hat{\alpha}_0$	0.0000792 (1.8874)
$\hat{\alpha}_1$	0.14999 (4.2839)
$\hat{\beta}_1$	0.5999 (7.607)
ARCH(1)	12.332

نلاحظ أن للنموذج MG-GARCH(1,1) معنوية إحصائية حيث نرفض الفرضية H_0 أي أن معاملات التوقع والتباين الشرطيين تختلف معنويا عن الصفر بنسبة معنوية 0.05 لأن قيم ستيودنت التي بين قوسين أكبر تماما من القيمة المجدولة للتوزيع الطبيعي 1.96. إضافة إلى ذلك، سلسلة مربعات البواقي تتميز باستقلالية تامة أن التباين الشرطي للبواقي متجانس باعتبار أن إحصائية ARCH-LM التي تساوي 1.205 أقل تماما من القيمة المجدولة لتوزيع χ^2 بدرجة حرية 1.

الشكل رقم (4): بواقي تقدير النموذج MG-GARCH(1,1)



للتأكد من ذلك، نطبق اختبار الاستقلالية BDS على بواقي التقدير المبينة في الشكل (4). النتائج تظهر في الجدول (6):

الجدول (6): نتائج اختبار BDS على بواقي MG-GARCH(1,1)

σ/ε		m
1	0.5	
-1.7504	-1.6928	2
-1.3558	-1.4647	3
-0.64331	-0.90981	4
-0.17411	-0.15938	5

من الملاحظ أن سلسلة البواقي تتميز باستقلالية تامة *i.i.d* حيث أن قيم BDS أقل تماما من القيمة الحرجة للتوزيع الطبيعي 1.96. بناء على ما سبق، يمكن القول أن مردودية مؤشر CAC40 تتميز ببنية مشوشة تتمثل في معادلة Mackey-Glass وعشوائية (خطأ من نوع GARCH).

2. الذاكرة الطويلة Long Memory Process: التفسير العشوائي للتقلبات

لقد بينت دراسات (1951) Hurst أن بعض السلاسل الزمنية تتميز ببنية ارتباط خاصة قريبة من عدم الاستقرارية. طور كل من (1968) .Mandelbrot and Van Ness و(1969) .Mandelbrot and Wallis نظرة Hurst ببناء حركة Brown الكسرية "Fractional Brownian Movements" ثم بعد ذلك التشويش ذات التوزيع الطبيعي الكسري "Fractional Gaussian Noise". تسمح هذه السيرورات بإحداث مركبات طويلة المدى لسلسلة زمنية. قد تتضمن سلسلة مستقرة مركبة الذاكرة الطويلة باعتبار أن تأثير القيمة الماضية على تلك الحالية تتناقص بوتيرة ضعيفة جدا. يسمى هذا السلوك بالارتباط طويل المدى أو "الصمود" وهذا يعني أن الاستجابة لصدمة عشوائية تعتبر كعودة نحو القيمة المتوسطة ولكن بسرعة جد ضعيفة.

النماذج الأساسية التي تسمح بتحديد الذاكرة الطويلة هي نماذج الانحدار الذاتي المتوسط المتحرك ذات التكامل الكسري ARFIMA " Autoregressive Fractionally Integrated Moving Average" التي اقترحها كل من (1980) .Granger and Joyeux و (1981) .Hosking.

1.2. تعريف السيرورة ARFIMA:

تعتبر هذه النماذج ترجمة في الزمن المتقطع لحركة Brown الكسرية. درجة التكامل ليست من الأعداد الصحيحة بل حقيقية. نقول أن السلسلة $\{Y_t, t = 1,2,...,T\}$ تخضع للسيرورة ARFIMA(p,d,q) إذا كان:

$$\phi_p(L)(1-L)^d Y_t = \theta_q(L)\varepsilon_t$$

حيث L معامل التأخير (أو التباطؤ)، $\phi_p(L)$ و $\theta_q(L)$ كثيرات الحدود المميزة من الدرجة p و q على الترتيب:

$$\phi_p(L) = 1 - \sum_{i=1}^{p} \phi_i L^i$$

$$\theta_q(L) = 1 - \sum_{j=1}^{q} \theta_j L^j$$

و ε_t تشويش أبيض ذو توقع رياضي معدوم وتباين ثابت σ^2.

يسمى $(1-L)^d$ بمعامل التكامل الكسري الذي يتفكك وفق صيغة النشر التالية:

$$(1-L)^d = 1 - dL + \frac{d(d-1)}{2!}L^2 + ... + (-1)^n \frac{d(d-1)...(d-n+1)}{n!}L^n + o(L^{n+1})$$

نقوم بإعطاء خصائص هذه السيرورة وذلك تبعا للقيم المختلفة لـ d.:

- إذا كان $d > -1/2$ و كل جذور كثير الحدود المميز $\theta_q(L)$ تقع خارج جذر الوحدة، فإن السيرورة $\{Y_t\}$ قابلة للقلب invertible.
- إذا كان $d < -1/2$ و كل جذور كثير الحدود المميز $\phi_p(L)$ تقع خارج جذر الوحدة، فإن السيرورة $\{Y_t\}$ مستقرة.
- إذا كان $-1/2 \leq d \leq 0$، فإن السيرورة $\{Y_t\}$ ضد الصمود anti-persistent.
- إذا كان $0 < d < 1/2$ ، فإن $\{Y_t\}$ سيرورة مستقرة بذاكرة طويلة (الاستقرارية طويلة المدى) يمكن استخدامها لنمذجة الصمود طويل المدى. في هذه الحالة، تتناقص دالة الارتباط الذاتي التي تعتبر موجبة بوتيرة بطيئة نحو الصفر (على شكل قطع زائد) عندما تكبر عدد الفجوات k.

تتضمن هذه السيرورة جزء ARMA التي تدرس بنية الارتباط قصير المدى و معامل التكامل الكسري الذي يشرح الحركة طويلة المدى. ترتبط الخصائص الأصلية للسيرورة

ARFIMA(p,d,q) بشكل المركبة طويلة المدى. يمكن دراسة هذه الخصائص بافتراض حالة السيرورة الكسرية ARFIMA(0,d,0). دالة الارتباط الذاتي $\rho(k)=\gamma(k)/\gamma(0)$ لهذه السيرورة تعرف كما يلي:

$$\rho(k)=\frac{\Gamma(k+d)\Gamma(1-d)}{\Gamma(k-d+1)\Gamma(d)}$$

والتي تكتب بالصيغة التقاربية التالية:

$$\rho(k)\sim\frac{\Gamma(1-d)}{\Gamma(d)}k^{2d-1}$$

دالة الكثافة الطيفية Spectral Density لهذه السيرورة هي:

$f(\omega)=\left[2\sin(\omega/2)\right]^{-2d}$ من أجل $\omega\in[0,\pi]$

والتي تكتب أيضا بالصيغة التقاربية كما يلي:

$$f(\omega)\equiv\omega^{-2d},\ |\omega|\to 0$$

في الحالة العامة، تُظهر دالة الكثافة الطيفية حالة شاذة (pick) عند الذبذبة Frequency 0 وتتناقص على شكل قطع زائد.

مثال 3:

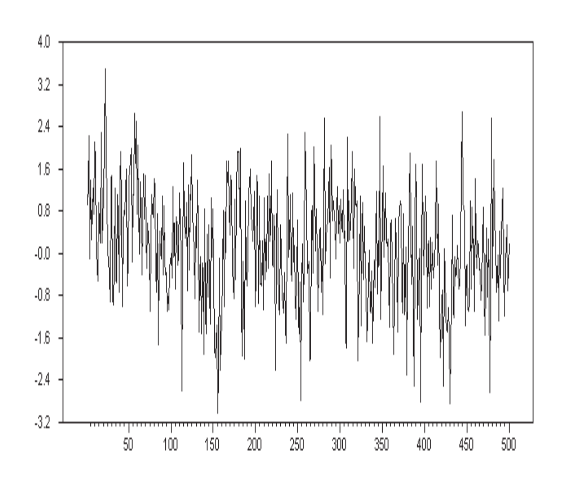
باستعمال تقنية المحاكاة، المطلوب بناء سلسلة زمنية تخضع للسيرورة ARFIMA(0, 0.23, 0):

$$(1-L)^{d}Y_{t}=\varepsilon_{t}$$

نستخدم برمجية RATS 5.04 في عملية المحاكاة مع $T=1000$:

```
ALLOCATE 1000
FREQUENCY 1 500
SOURCE(NOECHO) XGAMMA.SRC
SOURCE(NOECHO) ARFSIM.SRC
COMPUTE d = .23
COMPUTE  T = 500
CLEAR Y
@ARFSIM d T  Y
GRAPH 1
# Y
```

الشكل رقم (5): محاكاة سيرورة ARFIMA



يمكن أيضا استخدام برمجية GAUSS لهذا الغرض:

```
new;
library tsm,optmum,pgraph;
rndseed 123456;
d = 0.23;
Nobs = 1000;
{y,retcode} = RND_arfima(d,0,0,1,1000,Nobs,1);
t = seqa(1,1,Nobs);
graphset;
  _pdate = ""; _pnum = 2; _pltype = 6;
  title("Fractional processes");
  _plegctl = {2 7 1 1}; _plegstr = "d=-0.23\";
  xlabel("time");
  xy(t,y);
```

2.2. طرق تقدير معلم الذاكرة الطويلة:

1.2.2. الطرق الاستكشافية Heuristic methods:

تسمح هذه الطرق فقط بتقدير معلم التشابه الذاتي Auto-similarity H، يتعلق الأمر هنا بإحصائية R/S و أس Hurst.

تعرف إحصائية R/S كامتداد للمجاميع الجزئية للانحرافات بين السلسلة و متوسطها الحسابي مقسوما على انحرافها المعياري:

$$R/S = Q_T = \frac{1}{\sigma_Y}\left[\max_{1\le k\le T}\sum_{j=1}^{k}(Y_j-\overline{Y}_T)-\min_{1\le k\le T}\sum_{j=1}^{k}(Y_j-\overline{Y}_T)\right]$$

حيث σ_Y هو الانحراف المعياري للسلسلة، $\overline{Y}_T$ متوسطها و T حجم العينة. العبارة الأولى هي الحد الأقصى على k للمجاميع الجزئية لـ k. انحراف بين Y_j و متوسطها والعبارة الثانية هي الحد الأدنى على k للمجاميع الجزئية للانحراف.

لا ترتبط هذه الإحصائية بشكل التوزيع الهامشي للسلسلة، تسمح إذن بالكشف عن وجود بنية ارتباط طويل المدى في سلسلة زمنية معينة، إلا أن الإحصائية *R/S* لا تمثل اختبارا إحصائيا بمعنى الكلمة باعتبار أن التوزيع الاحتمالي غير معروف.

إن الميزة الأساسية لهذه الإحصائية أنها تعطي معاملا يسمى بأس Hurst الذي يسمح بتصنيف السلاسل الزمنية تبعا لنوع الارتباط. في هذه الحالة، يمكن تمثيل المعطيات بالعلاقة التالية:

$$R/S = \left(\frac{1}{2}T\right)^H$$

إذن:

$$H \approx \frac{\log(R/S)}{\log T}$$

يمكن تحديد مقياس للارتباط طويل المدى C_H المرتبط بأس Hurst. C_H يقيس الارتباط بين متوسط المشاهدات الماضية (الكبيرة) و متوسط المشاهدات المستقبلية (الكبيرة). الارتباط طويل المدى يعرف بالصيغة التالية:

$$C_H = 2^{2H-1} - 1$$

يمكن أن نصنف السلاسل الزمنية وفق القيمة المأخوذة من أس Hurst:

- إذا كان $H = 1/2$، فإن السيرورة لا تتميز بأي ارتباط طويل المدى، الارتباط C_H معدوم، مثلا التشويش الأبيض White Noise أو سيرورات الذاكرة القصيرة Short Memory processes.
- إذا كان $1/2 < H < 1$، فإن السلسلة تتميز بذاكرة طويلة، معاملات الارتباط الذاتي كلها موجبة تتناقص ببطء عندما تكبر الفجوة الزمنية (التباطؤ).

- إذا كان $0 < H < 1/2$، فإن السيرورة تعتبر في هذه الحالة ضد الصمود Anti-persistent: مراحل ارتفاع متبوعة بمراحل انخفاض. الارتباط C_H سالب، يتعلق الأمر هنا بضد الصمود طويل المدى.

لقد أشار Lo (1991) إلى هذه إحصائية R/S تمثل مشكلا حقيقيا يتمثل في شدة تأثرها بالارتباط قصير المدى للسلسلة. إذا كانت السلسلة الزمنية تتميز بذاكرة قصيرة، فتقدير أس Hurst باستعمال تحليل *R/S* متحيزا ولهذا السبب، اقترح Lo (1991) إحصائية *R/S* المعدلة والتي تأخذ الشكل التالي:

$$\widetilde{Q}_T = R/\hat{\sigma}_T(q) = \frac{1}{\hat{\sigma}_Y(q)}\left[\max_{1\le k\le T}\sum_{j=1}^{k}(Y_j - \overline{Y}_T) - \min_{1\le k\le T}\sum_{j=1}^{k}(Y_j - \overline{Y}_T)\right]$$

حيث:

$$\hat{\sigma}_T(q) = R/S_T = \frac{1}{T}\sum_{j=1}^{T}(Y_j - \overline{Y}_T)^2 + \frac{2}{T}\left[\sum_{j=1}^{q} w_j(q)\sum_{i=j+1}^{T}(Y_j - \overline{Y}_T)(Y_{i-j} - \overline{Y}_T)\right]$$

مع:

$$w_j(q) = 1 - \frac{j}{q+1}\ , \quad q < T$$

نلاحظ أن معاملات الارتباط الذاتي مبوبة (مرجحة) بدلالة الفجوات q والأوزان $w_j(q)$ تم وضعها من طرف Newey-West. (1987). بالمقابل، اقترح Andrews (1991) الاختيار التالي:

$$q = [k_T] = \left(\frac{3T}{2}\right)^{1/3}\left(\frac{2\hat{\rho}}{1-\hat{\rho}^2}\right)^{2/3}$$

$[k_T]$ يمثل الجزء الصحيح لـ k_T . و $\hat{\rho}$ تقدير الارتباط الذاتي من الدرجة الأولى. تعطى الأوزان $w_j(q)$ بالعلاقة $w_j = 1 - \left|\frac{j}{k_T}\right|$. يمكن القول أن التوزيع المحدود لإحصائية R/S المعدلة معروف ومرتبط بحركة Brown.

تطبيقيا، نمثل بيانيا هذه الإحصائية. إذا كانت السيرورة ذات ذاكرة طويلة، فهذا يعني أن النقاط يجب تتشتت عشوائيا حول خط مستقيم بميل $H > 1/2$ من أجل فجوات

زمنية كبيرة k_t وانطلاقا من الشكل البياني يمكن التفريق بين سيرورة ذات ذاكرة طويلة و ذات ذاكرة قصيرة.

مثال 4:

نقوم بتطبيق الطريقة المقترحة على مردودية سهم France Telecom باللوغاريتم خلال الفترة الممتدة بين 1998/01/05 و 2005/02/29. المطلوب تقدير أس Hurst و التمثيل البياني لإحصائية *R/S* المعدلة.

يحتوي برنامج RATS على ملف "source file" hurst.src الذي من خلاله يسمح بتقدير مقدر لمعلم Hurst و تمثيل إحصائية *R/S* بيانيا:

```
SOURCE(NOECHO) HURST.SRC
@HURST DLOGY /

Linear Regression - Estimation by Least Squares
Dependent Variable LOGRS
Usable Observations    540         Degrees of Freedom   538
Total Observations  1973631        Skipped/Missing   1973091
Centered R**2     0.936641          R Bar **2   0.936523
Uncentered R**2   0.997530          T x R**2     538.666
Mean of Dependent Variable          1.1976286025
Std Error of Dependent Variable     0.2414386612
Standard Error of Estimate          0.0608296081
Sum of Squared Residuals            1.9907297767
Regression F(1,538)               7953.2643
Significance Level of F           0.00000000
Durbin-Watson Statistic           0.055386

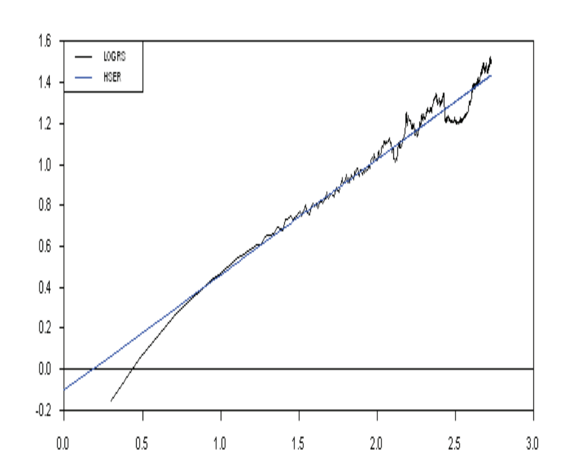
 Variable            Coeff          Std Error        T-Stat          Signif
*******************************************************************************
1.  Constant      -0.107111493     0.014862573      -7.20679       0.00000000
2.  LOGNOBS        0.565696829     0.006343238      89.18108       0.00000000

Hurst exponent =       0.56570
```

من خلال نتائج تقدير معلم Hurst، يتبين أن سلسلة مردودية سهم France Telecom تتميز بوجود ذاكرة طويلة لأن القيمة المقدرة لأس Hurst محصورة بين 0 و

½. الشكل (6) يؤكد هذه النتائج حيث أن إحصائية *R/S* معنوية باعتبار أن النقاط تتشتت عشوائيا حول خط مستقيم من أجل فجوات زمنية كبيرة.

الشكل رقم (6): تقدير إحصائية *R/S*



يمكن أيضا استعمال برمجية GAUSS لتقدير الإحصائيات المطلوبة وذلك بالاستعانة بالبرنامج التالي:

```
new;
library tsm,optmum,pgraph;
TSMset;
#include hurst.src;
load y[1,540]=ftelecom.dat;
{H,x} = HURST(y,2,5);
N = x[.,1];
Q = x[.,4];
V = x[.,5];
fit = x[.,6];
print ftos(h,"Hurst exponent: %lf",4,3);
```

نمثل بيانيا إحصائيتي R/S و R/S المعدل:

```
graphset;
  _pdate = ""; _pnum = 2;
  title("R/S statistic");
  xlabel("n"); ylabel("log(R/S)");
  LOGX(N,log(Q)~fit);
  title("V]n[ statistic");
  xlabel("n"); ylabel("V]n[");
  Bound = ones(rows(V),2).*(0.809~1.862);
  _pltype = 6~1~1;
  xy(N,V~Bound);
```

2.2.2. الطرق شبه المعلمية Semi-parametric methods:

اقترح Geweke and Porter-Hudak. (1983) طريقة تقدير شبه معلمية ترتكز على انحدار طيفي "Spectral Regression". بين هذان الباحثان أن معلم المعاينة لانحدار لوغاريتم الدالة الدورية periodogram على متغير مستقل تحديدي من أجل الذبذبات الأولى لـ Fourrier . $\omega_j = 2\pi j / T$ بطريقة المربعات الصغرى العادية يعتبر مقدرا متقاربا للمعامل d. ترتكز طريقة GPH على دالة الكثافة الطيفية المعطاة بالعلاقة التالية:

$$f(\omega_j) = \left|1 - e^{-i\omega_j}\right|^{-2d} f_\varepsilon(\omega_j),\ \omega \in [0, \pi]$$

حيث $f_\varepsilon(\omega_j) = \dfrac{\sigma^2 \left|\theta . e^{-\omega_j}\right|^2}{2\pi \left|\rho . e^{-iw_j}\right|^2}$ هي الكثافة الطيفية للسيرورة ARMA $(1-L)^d Y_t = \varepsilon_t$.

بإدخال اللوغاريتم على دالة الكثافة الطيفية، نحصل على:

$$\log f(\omega_j) = \log f_\varepsilon(0) - d \log\left|1 - e^{-i\omega_j}\right|^2 + \log \frac{f_\varepsilon(\omega_j)}{f_\varepsilon(0)}$$

لتكن $I_T(\omega_j)$ الدالة الدورية "periodogram":

$I_T(\omega_j) = \dfrac{1}{2\pi} \sum_{h=-T+1}^{T-1} \lambda(h \mid r) \gamma_h e^{-i\omega_j h}$، مع $\lambda(.)$ النافذة الطيفية و r معلم النافذة المختار بحيث يكون $0 < r(T) < T$. إذا اعتبرنا أن الذبذبات قريبة من الصفر، فالعبارة $\log\left(f_\varepsilon(\omega_j) / f_\varepsilon(0)\right)$ يمكن إهمالها. لدينا إذن:

حيث:

$$Y_j = \alpha + \beta X_j + \eta_j$$

$$Y_j = \log I_T(\omega_j)$$

$$X_j = \log\left|1 - e^{-i\omega_j}\right|^2$$

$$\eta_j = \log \frac{I_T(\omega_j)}{f(\omega_j)} - E(\eta)$$

$$\alpha = \log f_\varepsilon(0) + E(\eta)$$

$$\beta = -d$$

$g(T)$ دالة لـ T. متزايدة بحيث $m(T)/T \to 0$ لما: $T \to \infty$ $g(T) = T^m$ مع $0 < m < 1$. الأخطاء العشوائية η_j مستقلة ذاتيا و تخضع بصفة تقاربية لقانون χ^2 بدرجة حرية ν ترتبط بمعلم النافذة الطيفية.

مقدر معامل التكامل الكسري بطريقة المربعات الصغرى معطى بالصيغة التالية:

$$\hat{d}^{GPH} = \frac{\sum_{i=1}^{I}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^{I}\left(X_i - \bar{X}\right)^2}$$

إذا كان $-1/2 < d < 1/2$ و إذا وجدت متتالية m بحيث $(\log T)^2 / m \to 0$ لما $T \to \infty$، فإن مقدر OLS يتبع بصفة تقاربية التوزيع الطبيعي:

$$\hat{d}^{GPH} \to N\left(d, \frac{\operatorname{var}(\eta)}{\sum_{i=1}^{m}\left(X_i - \bar{X}\right)^2}\right) \text{ من أجل } T \to \infty$$

بين كل من Porter-Hudak (1990) و Crato et Lima (1994) أن الذبذبات T^m ينبغي اختيارها بحيث $I = T^m$ مع $m = 0.5, 0.6, 0.7, 0.8$.

مثال 5:

نقوم بتطبيق طريقة GPH لتقدير معلم الذاكرة الطويلة على السلسلة الشهرية لأسعار البترول باللوغاريتم (ذات الفروقات من الدرجة الأولى) log-differenced Oil Spot Prices خلال الفترة الممتدة بين شهر فبراير 1973 و ديسمبر 2008.

نأخذ في هذا المثال قيمتين من إحداثيات الدالة الدورية Periodogram. الهدف من اختيار قيمتين هو دراسة استقرارية المقدر عندما يتغير عدد إحداثيات الدالة الدورية.

يحتوي برنامج RATS على ملف "source file" GPH.SRC الذي من خلاله يسمح بتقدير مقدر لمعلم الذاكرة الطويلة بتقنية GPH:

من أجل $m = 0.8$:

```
SOURCE(NOECHO) GPH.SRC
@GPH(POWER=0.8) DLOIL /
```

من أجل $m = 0.5$:

```
SOURCE(NOECHO) GPH.SRC
@GPH(POWER=0.5) DLOIL /
```

الجدول (7): التقدير شبه المعلمي لمعامل التكامل الكسري باستخدام طريقة GPH

الإحداثيات	
$T^{0.8}$	$T^{0.5}$
0.274 (4.41)	0.35 (2.32)

القيم التي بين قوسين هي قيم ستيودنت

نلاحظ أن فرضية الذاكرة الطويلة مقبولة باستخدام اختبار GPH لأن قيم ستيودنت أكبر تماما من القيمة المجدولة للتوزيع الطبيعي 1.96 عند مستوى معنوية 5%، كما نلاحظ أيضا أن مقدر d محصور بين 0 و 0.5، مما يؤكد أن السلسلة تتميز بوجود ذاكرة طويلة.

نشير هنا أنه يمكن استعمال نوافذ طيفية أخرى مثل النافذة المثلثية، Bartlett، Daniell، Tuckey،... لتقدير معلم الذاكرة الطويلة d وباستخدام هذه النوافذ الطيفية، يعتبر المقدر أفضل بكثير من ذلك المحسوب بطريقة GPH[1].

بين Newbold and al (1991) أن مقدر GPH يعاني من مشكل أساسي يتمثل في وجود تحيز الدالة الدورية المقدرة estimated periodogram بصفة تقاربية. لهذا السبب، اقترح Robinson (1995b) طريقة سهلة تسمح فقط بتقدير معامل التكامل الكسري d بدون إعطاء أي معلومة تتعلق بتقدير المعالم الأخرى.

يعتبر Robinson (1995b) المقدر الجديد للمعلم H المقترح ذا توزيع طبيعي ولكن لا يفترض هذا المقدر أي صفة طبيعية للسيرورة.

لتحديد عبارة $\hat{H}$، لا بد أولا من إعطاء بعض التعريفات. لتكن $R(H)$ دالة معرفة بالعلاقة التالية:

$$R(H) = \log \hat{G}(H) - (2H-1)\frac{1}{m}\sum_{j=1}^{m}\log \lambda_i$$

و $\hat{G}(H)$ مقدرا لـ $G(H)$. المعطى بالصيغة:

$$\hat{G}(H) = \frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2H-1} I_j$$

حيث $m = \frac{1}{2}T$ من أجل T زوجي و $m = \frac{1}{2}(T+1)$ من أجل T فردي. في النظرية التقاربية، المعلم m يقترب ببطء نحو ∞.

يعطى المقدر $\hat{H}$ كما يلي:

$$\hat{H} = \arg\min_{H\in\theta} R(H)$$

مع $\theta = [\Delta_1, \Delta_2]$، Δ_1 و Δ_2 يتم اختيارهما بصفة كيفية بين 0 و 1 بحيث $0 < \Delta_1 < \Delta_2 < 1$.

1- Chikhi (2001), p. 69.

نشير هنا إلى أن $\hat{H}$ مقدر طبيعي Gaussian estimator لـ H_0 . في النموذج المعلمي التالي:

$$f(\lambda) = G_0 |\lambda|^{1-2H_0}, \lambda \in [-\pi, \pi]$$

مع G_0 و H_0 هي على الترتيب القيم الحقيقية لـ $G \in [0, \infty]$. و $H \in [0,1]$. لهذا المقدر سرعة تقارب تساوي $m^{1/2}$.

مثال 6:

باستعمال معطيات المثال السابق، نختبر فرضية الذاكرة الطويلة وفق تقنية Robinson بالاستعانة بتعليمة برمجية RATS:

```
ALLOCATE 455
FREQUENCY 1 455
SOURCE(NOECHO) RGSE.SRC
COMPUTE T = 455
SET X 1 T = DLOIL
@RGSE(POWER=.8) DLOIL 1 T DHAT DSE
```

نقوم بحساب القيمة المقدرة لـ d . والخطأ المعياري قبل حساب إحصائية ستيودنت:

```
DISPLAY 'DIFFERENCED d  :'#.### DHAT ####.### DSE
DIFFERENCED d  : 0.290   0.040
COMPUTE DHAT=0.220
COMPUTE DSE=0.040
COMPUTE TSTAT= DHAT/DSE
DISPLAY 'TSTAT='
TSTAT= 7.24
```

من خلال هذه النتائج، يمكن أن نؤكد أن للمعامل d معنوية إحصائية بنسبة معنوية 5% لأن إحصائية ستيودنت أكبر تماما من القيمة المجدولة 1.96 وهذا يعني أن هذه السلسلة تتميز بذاكرة طويلة.

إن التوزيع التقاربي لـ $\hat{d}$. لا يرتبط لا بمعالم الانحدار الذاتي و المتوسط المتحرك ولا بالتوزيع الاحتمالي للخطأ العشوائي للسيرورة ARFIMA. تعتبر إذن الطرق شبه المعلمية متحيزة في حالة وجود معالم AR وMA موجبة و كبيرة وأيضا سيكون هناك مبالغة أو

زيادة في تقدير الحركة طويلة المدى للسيرورة (over-estimation)، مما يجعلنا نفكر في استعمال طرق أخرى مثل طرق المعقولية العظمى عندما تكون التوزيعات الاحتمالية معروفة.

3.2.2. طرق المعقولية العظمى Maximum Likelihood Procedures:

يتعلق الأمر بتعظيم دالة المعقولية التامة أو التقريبية. يعتبر المقدر المتحصل عليه بطريقة المعقولية العظمى التامة المقترح من طرف Sowell (1992) الأفضل و الأقوى من بين المقدرات المقترحة في حالة وجود مركبة قصيرة المدى. يحتاج حسابه، عند كل تكرار خوارزمية التعظيم، إلى مصفوفة التباينات-التباينات المشتركة ومعكوسها.

أعطى Sowell (1992) دالة المعقولية العظمى غير الشرطية التامة لسلسلة زمنية مستقرة وطبيعية ذات تكامل كسري. لتكن y_t أي سلسلة زمنية و Y_T عينة مكونة من T مشاهدة حيث $Y_T = [y_1, y_2, ..., y_T]'$. Y_T تخضع للقانون الطبيعي بمتوسط معدوم ومصفوفة تباين-تباين مشترك Σ. تعرف دالة الكثافة كما يلي:

$$f(Y_T, \Sigma) = (2\pi)^{-T/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2} Y_T' \Sigma^{-1} Y_T\right)$$

حيث $|\Sigma|$ يمثل محدد مصفوفة التباين-التباين المشترك لـ Y_T.

يعطى لوغاريتم دالة المعقولية بالصيغة:

$$L(Y_T, \theta) = -\frac{T}{2}\log 2\pi - \frac{1}{2}\log|\Sigma(\theta)| - \frac{1}{2} Y_T' \Sigma^{-1}(\theta) Y_T$$

حيث θ هو شعاع المعالم الحقيقية غير المعروفة الذي يتضمن تباين الأخطاء، معلم التكامل الكسري و معالم الانحدار الذاتي و المتوسط المتحرك.

يتم الحصول على المقدر $\hat{\theta}$ لـ θ. وذلك بتعظيم لوغاريتم دالة المعقولية على شعاع المعالم θ أي:

$$\frac{\partial}{\partial\theta_j} L(Y_T, \theta) = -\frac{1}{2}\frac{\partial}{\partial\theta_j}\log|\Sigma(\theta)| - \frac{1}{2} Y_T' \frac{\partial}{\partial\theta_j}\Sigma^{-1}(\theta) Y_T$$

يعتبر المقدر $\hat{\theta}$ لـ θ حلا للمعادلة:

$$\frac{\partial}{\partial \theta_j} L(Y_T, \hat{\theta}) = 0$$

التوزيع التقاربي لمقدر طريقة المعقولية العظمى التامة يعطى بالصيغة $\hat{\theta} \xrightarrow{n.s} \theta$ لما $T \to \infty$ (*n.s*: بصفة أكيدة تقريبا) و $\sqrt{T}(\hat{\theta} - \theta)$ يؤول بتوزيع احتمالي إلى القانون الطبيعي بمتوسط معدوم و تباين-تباين مشترك $D^{-1}(\theta)$ مع:

$$D(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\frac{\partial}{\partial \theta_j} \log f(Y) \right) \left(\frac{\partial}{\partial \theta_j} \log f(Y) \right)' dY$$

اقترح Fox and Taqqu (1986) استعمال تقريب طيفي لدالة المعقولية العظمى التي تعطي مقدرات صحيحة و مضبوطة انطلاقا من عدد مشاهدات كاف. تعتبر هذه الطريقة أكثر فعالية ولكن نواجه صعوبة كبيرة في الميدان التطبيقي بسبب حساسيتها للشروط الابتدائية المعطاة للخوارزمية.

بصفة عامة، استخدام خوارزمية التعظيم تحتاج إلى اختيار قيم ابتدائية لمختلف معالم نموذج ARFIMA(p,d,q). لتحديد نموذج الذاكرة الطويلة، نتبع الخطوات التالية:

1. دراسة استقرارية السلسلة على المدى القصير وحساب السلسلة ذات الفروقات من الدرجة الأولى (في أغلب السلاسل الزمنية)
2. تقدير معامل التكامل الكسري d باستعمال تقنيات *R/S*، GPH و Robinson.
3. من أجل كل قيمة لـ d، نحسب السلسلة المحولة $(1-L)^{\hat{d}} Y_t$ ثم نقدر بطريقة Gauss-Newton النموذج ARMA(p,q) مع حساب تباين البواقي.
4. نختار كقيم ابتدائية المعالم المقدرة لنموذج ARMA الأمثل الذي يصغر تباين البواقي.
5. يتم تقدير معالم النموذج ARFIMA(p,d,q) وذلك بتعظيم دالة المعقولية على جميع المعالم وبإعطاء الدرجات (الرتب) القصوى $p = q = 3$.

6. نختار نموذج ARFIMA الأمثل بتصغير معياري AIC المصحح من طرف Hurvich and Tsai (1989)[1] و Schwarz[2].

مثال 7:

نطبق هذه الطرق على السلسلة اليومية لمردودية مؤشر FAZ الألماني باستعمال 6971 مشاهدة. يبين الجدول التالي نتائج تقدير نموذج ARFIMA(p,d,q):

الجدول (8): التقدير بطرق المعقولية العظمى

طريقة المعقولية العظمى التقريبية Fox-Taqqu		طريقة المعقولية العظمى التامة Sowell	
معيار SIC	معيار AICc	معيار SIC	معيار AICc
(0, 0.0642, 2) $t_d = 2.8576$ $ML = 22328.879$	(3, 0.0949, 3) $t_d = 2.5816$ $ML = 22334.158$	(2, 0.0511, 0) $t_d = 2.9007$ $ML = 22329.824$	(3, 0.0859, 3) $t_d = 2.4718$ $ML = 22334.929$

السطر الأول يظهر نموذج ARFIMA المختار وفق معيار AICc و SIC، t_d القيمة المحسوبة لإحصائية ستيودنت، أما السطر الثالث يمثل قيمة لوغاريتم المعقولية المعظمة.

من خلال هذا الجدول، نستنتج أن سلسلة المردودية تتميز بذاكرة طويلة أي قابلة للتنبؤ على المدى الطويل باعتبار أن إحصائيات ستيودنت أكبر تماما من القيمة المجدولة للتوزيع الطبيعي 1.96. نلاحظ أن هناك صمود طويل المدى وحركة سعر FAZ تظهر كنتيجة لصدمة خارجية مستدامة.

1- يعطى معيار AIC بالعلاقة التالية:

$$AICc = -2LV_{\max} + \frac{2kT}{T-(k+1)}$$

حيث $LV_{\max}$ هي لوغاريتم دالة المعقولية المعظمة، k عدد معالم النموذج و T حجم العينة.

[2] معيار Schwarz معرف كما يلي:

$$SIC = -(T-p)\ln\hat{\sigma}_\varepsilon^2 + k\ln(T-p)$$

حيث p هي درجة قسم AR و $\hat{\sigma}_\varepsilon^2 = \frac{1}{T-p}\sum_{t=p+1}^{T}\hat{\varepsilon}_t^2$، $\hat{\varepsilon}_t$ هي سلسلة البواقي.

في سيرورة ذات ذاكرة طويلة، معاملات الارتباط الذاتي تؤول إلى الصفر عندما تكبر الفجوة الزمنية ولكن تناقص هذه المعاملات يكون بوتيرة متمهلة أي على شكل قطع زائد، فتبعا لصدمة خارجية، تبتعد السلسلة الزمنية، خلال فترة زمنية طويلة نسبيا، عن مسارها السابق وتعود إلى قيمتها بعد فترة طويلة نسبيا.

3. التحليل غير المعلمي للسيرورات العشوائية بطريقة النواة

تستعمل هذا النوع من التقنيات في حالة ما إذا كانت بنية السيرورة غير الخطية غير معروفة، فيمكن تعميم النماذج غير الخطية إلى نماذج دالية لا تتضمن أي معلم. كان أول من أدخل هذه التقنية إلى السلاسل الزمنية هو كل من Bosq (1979)، Collomb (1980) و Robinson (1983).

في نموذج عشوائي، يوجد خصائص كثيرة لها تأثير مباشر على السلوك الحركي للنموذج مثل الصمود والتنبؤ، فهي ترتكز على اختيار المتغيرات المبطأة التي تبقى مفتوحة في النماذج المعلمية غير الخطية. إضافة إلى ذلك، نجد صعوبة كبيرة في تحديد النموذج المعلمي غير الخطي لأن هناك عدد لا يحصى من هذه النماذج، مما يؤدي إلى مشاكل أكثر تعقيدا في النمذجة، حيث أن الاختبارات غير قادرة على تحديد النموذج الملائم وخاصة إذا كانت السيرورة العشوائية تتميز بذاكرة قصيرة. كل الدراسات أثبتت قوة المقدرات غير المعلمية وتفوقها على تلك المعلمية الكلاسيكية، نذكر من بين هذه الدراسات Carbon and Délécroix (1993)، Gannoun (1991)، Rosa (1993) و Chikhi and Terraza (2002).

1.3. التقدير غير المعلمي لدالة الكثافة بطريقة النواة:

تلعب دالة التوزيع دورا مهما في دراسة القانون الاحتمالي لعينة معينة. غير أنها لا تسمح بالحصول على نتائج دقيقة حول هيكل هذا القانون. عوضا عن ذلك، عندما تكون لمتغيرات العينة دالة كثافة g، فهذه الأخيرة تعطي معلومات أكثر دقة على القانون

(تشتت، منوال..الخ) وهذا يفسر لماذا تقدير g يعتبر مشكلا أساسيا في الطرق غير المعلمية. يكمن الهدف الأساسي إذن في تقدير دالة كثافة المتغيرات $Y_t, t \in Z$ التي تشكل سيرورة مستقرة. نعتبر هنا مقدر Parzen-Rosenblatt:

$$\hat{g}(y) = \frac{1}{T.h_n^k} \sum_{t=1}^{T} K\left(\frac{y - Y_t}{h_n}\right)$$

حيث T حجم العينة، h_n تمثل متتالية موجبة تؤول إلى الصفر لما حجم العينة يؤول إلى ∞ و $K(.)$ دالة النواة.

نستخدم بصفة عامة مدرج تكراري متحرك، أي نأخذ بعين الاعتبار المشاهدات الواقعة في مجال طوله h_n ومركز حول y من الشكل $]y - h_n/2, y + h_n/2[$، عدد النقاط التي تقع في هذا المجال تحقق:

$$-\frac{1}{2} \leq \frac{y - Y_t}{h_n} \leq \frac{1}{2}$$

يمكن اختيار النواة بحيث يكون الوزن الموجود على المشاهدات Y_t بمقدار جد ضئيل بالمقارنة مع المشاهدات البعيدة عن y. يمكن أن نعرف أنواعا كثيرة من مقدرات الكثافة باختيار K منها: النواة الطبيعية، Epanechnikov، الثنائية....الخ. يعتبر مقدر النواة ممهدا بالنسبة للمدرج التكراري الذي هو ثابت بقطعة ومتقاربا بصفة سريعة وهذا ما يؤدي بنا إلى وضع فرضيات إضافية على النواة. في هذه الحالة، نفرض أنها موجبة تماما، متناظرة، محدودة وقابلة للتكامل، تحقق الخصائص التالية:

1) $\|y\|^k |K(y)| \to 0$ عندما $\|y\| \to \infty$

2) $\int_{R^k} \|y\|^2 K(y)dy$ حيث $\|y\|$ هو الشكل الاقليدي على R^k

3) $\int K(y)dy = 1$

تجدر الإشارة إلى أن سرعة تقارب المقدرات غير المعلمية هي من الرتبة $T^{-1/(k+4)}$ واختيار النافذة Bandwidth مهم للغاية لأنها تحدد درجة تمهيد المقدر.

مثال 8:

قررنا التأكد من صحة هذه النتائج على سلسلة مصطنعة عن طريق محاكاة تخضع للتوزيع الطبيعي المختزل المركز. بالاستعانة ببرمجية GAUSS 3.2 – TSM Module –، نقدر دالة الكثافة بطريقة نواة Epanechnikov، يتغير طول السلسلة من 50 إلى 100000.

```
new;
library tsm,optmum,pgraph;
TSMset;
rndseed 123;
y = rndn(T,1);  /* T = 50,150,.....,50000,100000*/
{x,dens,F,retcode} = Kernel(y);
graphset;
 _pdate = ""; _pnum = 2;
 title("Probability density function");
 _plegstr = "estimated\0true";
 _plegctl = 1;
 xy(x,dens~pdfn(x));
```

تظهر النتائج في الجدول التالي:

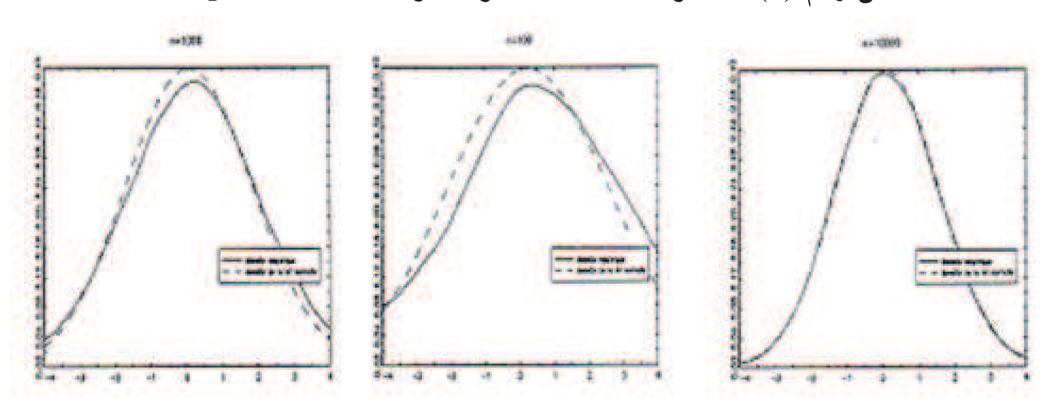
الجدول (9): نتائج محاكاة معلم التمهيد (النافذة)

النافذة h_n	T
0.4889	50
0.3854	150
0.3151	500
0.2738	1000
0.1918	5000
0.1743	8000
0.1667	10000
0.1215	50000
0.1059	100000

يتضح جليا من خلال الجدول المبين أعلاه أنه إذا كان كبر عدد المعطيات المستعملة، تنخفض قيمة النافذة، فنافذة كبيرة جدا تؤدي إلى تمهيد مبالغ فيه Over-Smoothing. فضلا عن ذلك، إذا زادت h_n، التباين يتناقص ولكن مقدار التحيز يزداد. في هذه الحالة، نختار نافذة تُوازن بصفة تقاربية بين التباين و التحيز. اقترح Deheuvels (1977) لتقدير

دالة الكثافة الاختيار التالي $h_n = \hat{\sigma}_n T^{-1/(k+4)}$ (هذا الاختيار لا يأخذ بعين الاعتبار شروط المزج Mixing Conditions) حيث $\hat{\sigma}$ يمثل مقدار الانحراف المعياري للسلسلة.

الشكل رقم (7): تقدير دالة الكثافة بطريقة نواة Epanechnikov

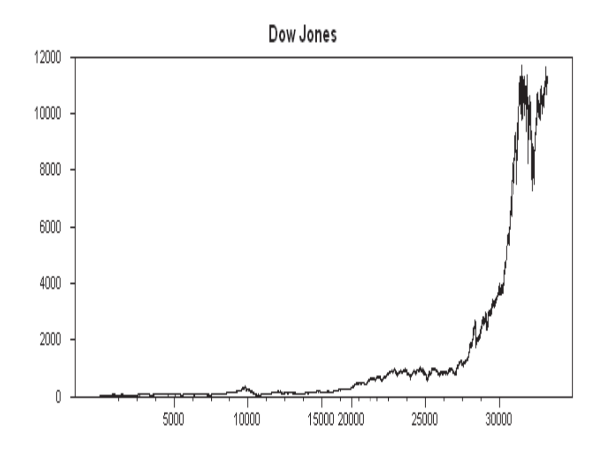


لقد رأينا أنه إذا كانت النافذة كبيرة جدا، فهذا يؤدي إلى سرعة تقارب غير جيدة. عندما يكون عدد مشاهدات السلسلة يساوي 100، نلاحظ بسهولة أن دالة الكثافة المقدرة لا تقترب بصفة كلية من تلك النظرية. كلما كان حجم العينة كبيرا، كلما كانت سرعة التقارب جيدة، هذا يعني أنه إذا كانت T صغيرة جدا والنافذة كبيرة، فسرعة التقارب غير مثلى. في هذه الحالة يكبر مقدار التحيز (أنظر الشكل (7)).

مثال 9:

نقوم بتقدير دالة الكثافة لسلسلة لوغاريتم Dow Jones ذات الفروقات من الدرجة الأولى بطريقة النواة في الفترة الممتدة بين 1896/05/26 و 2006/08/17. يُظهر الشكل البياني التالي تطور سلسلة مؤشر Dow Jones:

الشكل رقم (8): سلسلة مؤشر Dow Jones

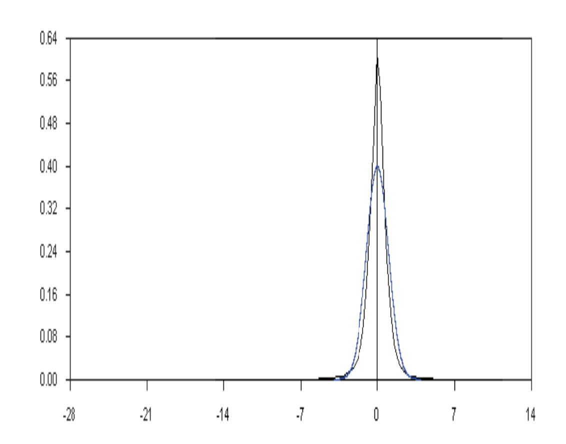


لتقدير دالة الكثافة بطريقة النواة نستخدم برمجية RATS 5.04 وفق التعليمة التالية:

```
LOG DJ / LOGDJ
DIFF LOGDJ / DLOGDJ
SOURCE(NOECHO) KERNEL.SRC
@KERNEL(KERNEL=EPANECHNIKOV,STYLE=LINES,NGRAPH)  DLOGDJ /
```

الشكل رقم (9): تقدير غير معلمي لدالة الكثافة

بنواة Epanechnikov $(h_n = 0.1345)$



2.3. سيرورة الانحدار الذاتي غير الخطي:

يمكن كتابة سيرورة الانحدار الذاتي غير الخطي $\{Y_t\}_{t\geq 0}$ مع خطأ ARCH (-NAR ARCH) رياضيا كما يلي:

$$Y_t = f(X_t) + \sigma^{1/2}(X_t)\varepsilon_t$$

حيث $X_t = (Y_{t-i_1},.......,Y_{t-i_k})'$ منقول شعاع للمتغيرات المبطأة $i_1 < .... < i_k$ و $\varepsilon_i \sim i.i.d(0,1)$. لنفرض أن كل التباطؤات (Lags) مهمة لتحديد دالة المتغيرات المستقلة $f(.)$.

الفرضيات المتعلقة بهذه السيرورة معطاة من طرف Doukhan (1994):

(H.1) من أجل كل عدد صحيح $M \geq i_k$، السيرورة $X_{M,t} = (Y_{t-1},.....,Y_{t-M})^{\mathrm{T}}$ مستقرة تماما و يمتاز بخاصية المزج القوي $\beta -$ mixing مع $\beta(T) \leq c_0 T^{-(2+\delta)/\delta}, \delta > 0, c_0 > 0$ حيث:

$$\beta(T) = E\sup\left\{\left|P(A|\Re_M^k) - P(A)\right| : A \in \Re_{T+k}^{\infty}\right\}$$

حيث $\Re_t^{t'}$ هي σ-جبر لـ . $X_{M,t}, X_{M,t+1},......, X_{M,t'}$.

(H.2) التوزيع المستقر للسيرورة $X_{M,t}$ لها دالة كثافة $g_M(x_M), x_M \in R^M$ مستمرة. إذا استعملنا مقدر Nadaraya-Watson، فيجب أن تكون الدالة $g_M(.)$ قابلة للتفاضل بصفة (H.3) الدالة $f(.)$ قابلة للتفاضل مرتين بصفة مستمرة و الدالة $\sigma(.)$ موجبة ومستمرة.

(H.4) العزوم الأربعة الأولى لـ . $\{\varepsilon_t\}_{t\geq i_k}$ محدودة.

(H.4) $K: R \to R$ نواة موجبة ومتناظرة و $h = h_n$ عدد موجب مع $h \to 0, Th^k \to \infty$ عندما تكون $T \to \infty$.

يُلاحظ أن: $\|K\|_2^2 = \int K^2(u)du, \sigma_K^2 = \int K(u)u^2 du$. من أجل $x \in R^k$، لدينا:

$$K_h(x) = \frac{1}{h^k}\prod_{j=1}^{k} K(\frac{x_j}{h})$$

يمكن تقدير دالة التوقع الشرطي $f(.)$ والتباين (أو التطاير) الشرطي $\sigma(.)$ باستخدام تقنيتين: مقدر Nadaraya-Watson $\hat{f}_1(x)$ و المقدر الخطي المحلي (أو الموضعي) Local Linear Estimator $\hat{f}_2(x)$.

ليكن:

$$X_t = (Y_{t-i_1}, Y_{t-i_2}, ..., Y_{t-i_k})' = (X_{t_1}, X_{t_2}, ..., X_{t_k})'$$

و

$$Y = (Y_{i_k}, Y_{i_{k+1}}, ..., Y_n)'$$

من أجل $x \in R^k$ ، يمكن كتابة المقدرين رياضيا على الشكل التالي:

$$\hat{f}_1(x) = (Z_1'WZ_1)^{-1} Z_1'WZ$$
$$\hat{f}_2(x) = e'(Z_2'WZ_2)^{-1} Z_2'WZ$$

حيث:

$$Z_1 = (1, ..., 1)'_{1\times(T-i_k+1)}$$

$$Z_2 = \begin{pmatrix} 1 & . & . & . & 1 \\ X_{i_k} - x & . & . & . & X_T - x \end{pmatrix}'$$

$$e = \begin{pmatrix} 1 & 0_{1\times k} \end{pmatrix}'$$

$$W = diag\left\{\frac{1}{T'} K_h(X_i - x)\right\}_{i=i_k}^{T}$$

مع: $T' = T - i_k + 1$

اقتُرح التقارب The convergence و التوزيع الطبيعي التقاربي Asymptotic Normality لمقدر Nadaraya-Watson للنماذج NAR-ARCH من طرف Masry and Tjostheim (1995)، لدينا:

$$\sqrt{T'h^k}\left\{\hat{f}_1(x) - f(x) - h^2(b_c(x) + b_t(x))\right\} \xrightarrow{D} N(0, V(x))$$
$$\sqrt{T'h^k}\left\{\hat{f}_2(x) - f(x) - h^2 b_t(x)\right\} \xrightarrow{D} N(0, V(x))$$

وهذا عندما يكون $T \to \infty$ مع التحيز التقاربي:

$$b_c(x) = \frac{\sigma_K^2}{2}\left\{\frac{2\nabla f(x)^{\mathrm{T}}\nabla\mu(x)}{g(x)}\right\}$$

$$b_t(x) = \frac{\sigma_K^2}{2}Tr\left[\nabla^2 f(x)\right]$$

والتباين التقاربي:

$$V(x) = \frac{\sigma(x)}{g(x)}\|K\|_2^{2k}$$

حيث $\|K\|_2^2 = \int K^2(x_j)dx_j$ و $\sigma_K^2 = \int x_j^2 K(x_j)dx_j$.

لهذين المقدرين نفس التباين التقاربي. ترتبط هذه الأخيرة بدالة التباين الشرطي $\sigma(x)$ و دالة الكثافة $g(x)$. تعتمد نسبة تقارب المقدرات Convergence rate of estimators على بعد مصفوفة المتغيرات المستقلة وهذان المقدران يستعملان لبناء فترات ثقة لـ . $\hat{f}_i(x)$, $i = 1,2$.

3.3. تحديد وتقدير سيرورة غير معلمية و معيار *CAFPE*:

يجب تحديد النافذة h وشعاع التباطؤات الزمنية (الذي يسمى بمعاملات Markov) $i_1,\ldots\ldots,i_k$، فيُعتبر اختيار النافذة مهما لأنها تحدد مستوى تمهيد الدالة المقدرة. لهذا نستعمل عادة طريقة Cross-validation حيث أن خصائصها التقاربية تم دراستها من طرف (1994) Vieu و (1994) Yao and Tong، إلا أن هناك طريقة بديلة ترتكز على معيار *FPE* "Final Prediction Error" واقترحت من طرف Tjostheim and Auestad (1994) لدراسة المقدر غير المعلمي لـ . Nadaraya-Watson. يمكن القول أن المعيار التقاربي يعطي أحسن سرعة تقارب أفضل من معيار Cross-validation. استخدم مقدر Nadaraya-Watson في عدة طرق لتحديد معاملات Markov. لهذا المقدر نسبة تقارب ضعيفة جدا عندما تكون دالة كثافة المتغيرات المبطأة، خاصة في السيرورات غير الخطية، غير ممهدة بشكل كاف. عكس ذلك، المقدر الخطي المحلي Local Linear Estimator يحتاج فقط إلى استمرارية دالة الكثافة لكي تكون نسبة التقارب مثلى.

لتكن $\{\tilde{Y}_t\}$ سلاسل لها نفس توزيع السلسلة $\{Y_t\}$ ولكن مستقلة عن هذه الأخيرة.
يعرف معيار FPE لمقدر $\hat{f}$ لـ f . بالعلاقة التالية:

$$FPE(\hat{f}) = \lim_{t\to\infty} E\left[\left\{\tilde{Y}_t - \hat{f}(\tilde{X}_t)\right\}^2 w(\tilde{X}_{M,t})\right]$$

حيث التوقع (الأمل) يعطى على كل المتغيرات $Y_0, Y_1, ..., Y_n, \tilde{Y}_0, \tilde{Y}_1, ..., \tilde{Y}_n$.

(H.6) $w(.)$ مستمرة وموجبة، نفترض أيضا أن $g(x_M) > 0$ من أجل supp(w) $x_M \in$.(supp: support)

إذا كانت $\{\tilde{Y}_t\}$ سيرورة انحدار ذاتي غير خطية بصفة تقاربية ومستقرة و $\hat{f}$ مقدرا غير معلمي، فإن المعيار FPE غير خطي. لدينا:

$$FPE_a(h) = AFPE_a(h) + o(h^4 + \frac{1}{T' h^k})$$

حيث:

$$AFPE_a(h) = A + b(h)B + c(h)C_a$$

مع:

$$A = \int \sigma(x) w(x_M) g(x_M) dx_M$$
$$B = \int \sigma(x) w(x_M) g(x_M) / g(x) dx_M$$
$$C_a = \int r_a^2(x) w(x_M) g(x_M) dx_M$$
$$r_1(x) = Tr\{\nabla^2 f(x)\} + 2\nabla^T g(x) \nabla f(x) / g(x)$$
$$r_2(x) = Tr\{\nabla^2 f(x)\}$$

مع العلم أن:

$$b(h) = \|K\|_2^{2k} T'^{-1} h^{-k}, c(h) = \sigma_K^4 h^4 / 4$$

يشير $a = 1,2$ إلى مقدر Nadaraya-Watson و المقدر الخطي المحلي (الموضعي) Local Linear Estimator. تمثل العبارة الأولى A توقع تباين التشويش الأبيض للسيرورات $\{Y_t\}$ و $\{\tilde{Y}_t\}$. العبارة الثانية B توقع تباين المقدر غير المعلمي $\hat{f}_a$ ، أما العبارة الثالثة C تمثل توقع تحيزه.

نذكر أن هذين المقدرين مختلفان من حيث مقدار التحيز وبالتالي المعياران التقاربيان مختلفان أيضا، القيم المتوقعة مستقلة عن النافذة h وتتحدد بخصائص السيرورة التي تُمثل

المعطيات Data Generator Process (DGP). يصبح التوقع الرياضي للتباين ومقدار التحيز مهملين و $AFPE(h)$ يؤول إلى قيمته الصغرى التقاربية $E(\sigma(x))$ لما $h \to 0$ و $Th^k \to \infty$ من أجل $T \to \infty$. غير أنه من أجل عينة معطاة T، نحصل على عبارة تحدد نافذة مثلى $h_{a,opt}$ بتصغير مجموع العبارتين الثانية والثالثة:

$$h_{a,opt} = \left\{k \|K\|_2^{2k} Bn'^{-1} C_a^{-1} \sigma_K^{-4}\right\}^{1/(k+4)}$$

و معيار $AFPE$ "Asymptotic Final Prediction Error" المصغر هو:

$$AFPE_{a,opt} = A + (k^{-k/(k+4)} + \frac{1}{4} k^{4/(k+4)}) \left\{\|K\|_2^{8k} B^4 T'^{-1} C_a^k \sigma_K^{4k}\right\}^{1/(k+4)}$$

باعتبار أن القيمتين A و B مجهولتان، فلا بد من تقديرهما ليتسنى لنا في الأخير تقدير المعيار $AFPE$ والنافذة المثلى. لدينا:

$$\hat{A}_a = T'^{-1} \sum_{i=i_k}^{n} \left\{Y_i - \hat{f}_a(X_i)\right\}^2 w(X_{M,i})$$

$$\hat{B}_a = T'^{-1} \sum_{i=i_k}^{n} \left\{Y_i - \hat{f}_a(X_i)\right\}^2 w(X_{M,i}) / \hat{g}(X_i)$$

يستعمل المقدران $\hat{f}_a$ نافذات من الرتبة $T'^{-1/(k+4)}$ و $\hat{g}(X_i)$ مقدر دالة الكثافة بطريقة النواة. تحت الفرضيات (H.1)-(H.6)، من أجل $a = 1,2$، لما $T \to \infty$:

$$\hat{A}_a = A + \left\{\|K\|_2^{2k} - 2K(0)^k\right\} T'^{-1} h^{-k} B + C_a \sigma_K^4 h^4 / 4 + o\left\{h^4 + T'^{-1} h^{-k}\right\} + O_P\left\{T'^{-1/2}\right\}$$

تجدر الإشارة الى أن المقدر غير المعلمي $\hat{A}_a$ تقترب نحو A وتقاربه من الرتبة $\sqrt{T}$ إذا كان $k \leq 4$ مع نافذة h من الشكل $c.T^{-1/(k+4)}$. في كل الحالات، ستكون العبارة الثانية و الثالثة من الرتبة $O\left\{T'^{-1/2}\right\}$، مقدر FPE التقاربي غير المعلمي يعطى بالعلاقة:

$$AFPE_a = \hat{A}_a + 2K(0)^k T'^{-1} h_{a,opt}^{-k} \hat{B}_a$$

$\hat{A}_a$ مقدر باستخدام نافذة مثلى $h_{a,opt}$، أما $\hat{B}_a$ مقدر باستخدام أي نافذة من الرتبة $T'^{-1/(k+4)}$.

من جهة أخرى، يمكن إضافة معامل تصحيح للمعيار $AFPE_a$ وذلك حتى لا يكون هناك مبالغة أو زيادة في تقدير معاملات Markov " Over-estimation of Markov coefficient". المعيار $AFPE_a$ المعدل (أو المصحح) والذي نرمز له بـ . $CAFPE_a$ "Corrected Asymptotic Final Prediction Error" يعرف كما يلي:

$$CAFPE_a = AFPE_a \left(1 + \frac{k}{T^{\frac{4}{k+4}}} \right)$$

حيث أن التصحيح يرتبط بالمعاملات k وعدد المشاهدات T. يتطلب إذن اختيار مجموعة جزئية $\{\hat{i}_1,......,\hat{i}_{\hat{k}}\}$ تحت الفرضيات (H.1)-(H.6)، بحيث:

$$P\{\hat{k} = k, \hat{i}_l = i, l = 1,......,k\} \xrightarrow[T \to \infty]{} 1$$

لتقدير النموذج غير المعلمي NAR-ARCH باختيار معاملات Markov والنافذة، نستعمل معيار $CAFPE$. لنعتبر عدد التباطؤات القصوى d والتباطؤ الأكبر هو p، ينبغي إذن فحص $1 + \sum_{j=1}^{d} \binom{k_{\max}}{j}$ توفيقة مختلفة لمعاملات Markov (التباطؤات) في طريقة البحث. الهدف الأساسي هو اختيار تباطؤات تجعل المعيار $CAFPE$ أصغر ما يمكن. في الخطوة الأولى، التباطؤ المهم هنا هو $\hat{i}_1$ يُحدد باختيار i_1 بحيث يكون المعيار أصغر ما يمكن للنموذج ذي تخلف زمني واحد. إذا كان المعيار مُقارنا بتباين السلسلة، ليس هناك أي تباطؤ معنوي و إلا نبحث عن $\hat{i}_2$ في الخطوة الموالية بتثبيت $\hat{i}_1$ وباختيار i_2 بحيث يكون العيار مُصغرا للنموذج ذي تخلفين زمنيين. كذلك، نُدخل $\hat{i}_2$ وتستمر الطريقة إذا كان المعيار على تخلفين زمنيين أصغر من قيمة المعيار على تخلف زمني واحد. نستعمل نفس الطريقة لحساب $\hat{i}_3 ,....,\hat{i}_d$ ..

عند حساب المعيار في كل خطوة، نحسب النافذة التي تُحدد درجة تمهيد المقدر. لحساب البواقي ودالة الكثافة التي تُستخدم لتقدير التوقع الرياضي لتباين المقدر غير المعلمي $\hat{f}_a$ ، نستعمل نافذة Silverman المعرفة رياضيا كما يلي:

$$h_S = \text{var}(Y_t)\left(\frac{4}{k+2}\right)^{\frac{1}{k+4}} T^{-\frac{1}{k+4}}$$

لتقدير النافذة المثلى، نستخدم طريقة البحث التشابكي "Grid Search" من أجل $a=1,2$ $CAFPE_a$, حيث القيم الممكنة تقع في المجال $[0,2h_S, 2h_S]$ في 24 خطوة (24 steps). نختار النافذة المثلى التي تُصغر المعيار $CAFPE_a$ والتي تنتمي إلى هذا المجال.

لتقدير سلسلة زمنية معينة، نفضل استعمال النوة الطبيعية لأن استخدام أنواع أخرى من النواة يتطلب استعمال نافذة أكبر. تعرف دالة الوزن $w(.)$ كما يلي:

$$w(X_t) = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} \text{ if } \begin{cases} \hat{g}(X_t) \geq c \\ \hat{g}(X_t) < c \end{cases}$$

حيث تُحدد الثابتة c بحيث يتم استبعاد T'' مشاهدة و T'' يمثل أكبر عدد صحيح بعد القيمة $0,05T'$.

مثال 10:

نطبق طريقة التقدير غير المعلمي على السلسلة اليومية لمردودية مؤشر CAC40 خلال الفترة الممتدة بين 1987/09/07 و 1999/05/28.

إن سلسلة مردودية مؤشر CAC40 مستقرة حيث أثبت chikhi (2001) أنها لا تحتوي على أي جذر وحدوي وعليه نقوم بدراسة خصائصها الإحصائية قبل المضي في تطبيق الطريقة غير المعلمية. الجدول التالي يعطي الخصائص الإحصائية لهذه السلسلة:

الجدول (10): الخصائص الإحصائية لسلسلة مردودية مؤشر CAC40

إحصائية Jarque-Bera	معامل Kurtosis	معامل Skewness
5017.611	9.210	-0.447

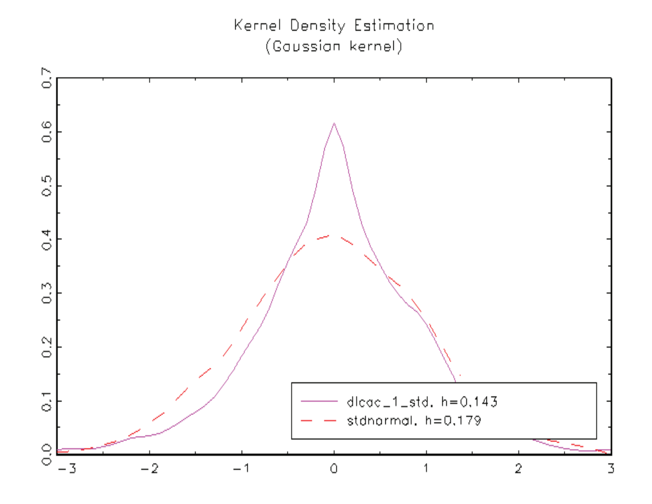
من خلال هذا الجدول يتبين أن هذه السلسلة تتميز بتوزيع غير طبيعي باعتبار أن إحصائية Jarque-Bera التي تساوي 5017.611 أكبر تماما من القيمة المجدولة لتوزيع χ^2

بدرجة حرية 2، كما أن هذا التوزيع يتسم بعدم التناظر و ملتو نحو اليسار (معامل Skewness سالب)، فقد يكون عدم التناظر دليلا على عدم خطية هذه السلسلة مثلا وجود أثر ARCH الذي نصادفه كثيرا في السلاسل المالية (معامل Kurtosis أكبر من 3) و أيضا وجود بنية مشوشة كما رأينا في المثال 2.

يمكن التأكد من نتيجة اختبار Jarque-Bera بتقدير دالة الكثافة بطريقة النواة ومقارنتها بدالة كثافة التوزيع الطبيعي (أنظر الشكل (10)):

الشكل رقم (10):

تقدير غير معلمي لدالة كثافة مردودية مؤشر CAC40 بطريقة النواة



من الملاحظ أن دالة الكثافة المقدرة لا تقترب كثيرا من تلك النظرية وهذا يعني أن التوزيع غير طبيعي. كما نستنتج أيضا من خلال نتائج اختبار BDS في المثال 2 أن سلسلة مردودية مؤشر CAC40 قابلة للتنبؤ على المدى القصير حيث رفضنا فرضية الاستقلالية أي أن هناك بنية ارتباط غير خطي بين المشاهدات، يمكن التأكد من ذلك باستعمال اختبار الاستقلالية لـ . Mizrach. النتائج تظهر في الجدول الموالي:

الجدول (11): نتائج اختبار Mizrach

إحصائيات Mizrach $T \leq 1000$	m
3.495	1
4.072	2
4.951	3
3.537	4
2.969	5
2.566	6
2.381	7
2.51	8
2.648	9
2.326	10

الجدول (11) يُظهر بنية ارتباط قوية بين مشاهدات السلسلة، حيث نرفض فرضية العدم للاستقلالية باعتبار أن القيم المحسوبة أكبر تماما من القيمة المجدولة للتوزيع الطبيعي 1.96.

من خلال هذه الاختبارات، يتبين أن السوق الفرنسي غير كفء بصفة ضعيفة حيث أظهرت هذه الاختبارات قابلية هذه السلسلة للتنبؤ على المدى القصير ولكنها لا تسمح بالكشف عن بنية ارتباط طويل المدى وعليه لا بد من اختبار قابلية التنبؤ على المدى الطويل وذلك باختبار الذاكرة الطويلة للسلسلة بتقدير معامل التكامل الكسري بطريقة GPH شبه المعلمية وباستخدام مجموعة من النافذات الطيفية.

الجدول (12): نتائج تقدير معامل الذاكرة الطويلة لسلسلة المردودية:

تقدير نموذج ARFIMA(0,d,0)

النافذة	الإحداثيات				
	$T^{0.65}$	$T^{0.6}$	$T^{0.55}$	$T^{0.5}$	$T^{0.45}$
GPH	0.042 (0.850)	0.059 (0.964)	0.119 (1.539)	0.089 (0.914)	0.088 (0.717)
Rectangular .	0.039 (0.716)	0.038 (0.557)	0.049 (0.576)	0.016 (0.155)	0.041 (0.305)
Bartlett	0.046 (1.454)	0.051 (1.299)	0.049 (0.576)	0.042 (0.687)	0.094 (1.21)
Daniell	0.046 (1.194)	0.051 (1.07)	0.068 (1.385)	0.042 (0.561)	0.095 (0.993)
Tukey	0.043 (1.269)	0.047 (1.112)	0.063 (1.178)	0.035 (0.517)	0.078 (0.92)
Parzen	0.045 (1.597)	0.05 (1.413)	0.069 (1.554)	0.042 (0.760)	0.092 (1.313)
B-priest	0.041 (0.984)	0.043 (0.825)	0.057 (0.867)	0.027 (0.325)	0.062 (0.59)

إن سلسلة مردودية مؤشر CAC40 لا تتضمن إلا مركبة قصيرة المدى أي نرفض فرضية الذاكرة الطويلة باعتبار أن قيم ستيودنت التي بين قوسين أقل تماما من القيمة المجدولة للتوزيع الطبيعي 1.96 وذلك في كل النافذات الطيفية، فوجود حركية قصيرة المدى يشير إلى أن الوكلاء لا يستطيعون التوقع بالمردودية في فترة طويلة. تظهر حركة الأسعار كنتيجة لصدمة عابرة وليست مستدامة.

بناء على ما تقدم، نعمم النماذج غير الخطية المعلمية إلى صيغة غير معلمية بحيث يتم تقدير مردودية CAC40 بطريقة النواة. في هذا المثال، العدد الأقصى للتباطؤ هو $d = 3$ والتباطؤ الأكبر هو $k_{\max} = 8$. لكي نحدد النموذج المناسب للسلسلة، نقوم بتقدير معاملات Markov والنافذة المثلى وذلك بتصغير معيار *CAFPE*. الطريقة المستعملة هي المقدر الخطي الموضعي Local Linear Estimator. من أجل هذا نستخدم برمجية

GAUSS بالاستعانة بثلاث ملفات "Library Files": density.dll، locling.dll و loccubg.dll. مثلا يجب تحديد بدقة طريق هذه ملفات المكتبة كما يلي:

```
library optmum,multband, pgraph;
   dlibrary c:\gauss\dlib\locling.dll,
   c:\gauss\dlib\density.dll,
   c:\gauss\dlib\loccubg.dll;
```

النتائج مبينة في الجدول التالي:

الجدول (13): نتائج تقدير نموذج NAR و ARCH غير المعلمي

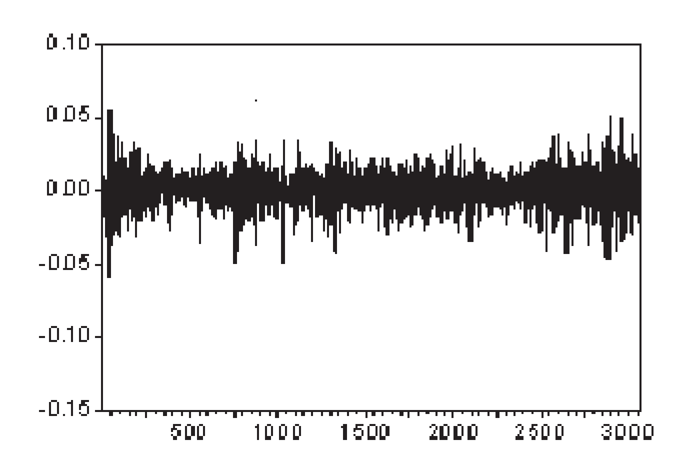
النموذج	الفجوات الزمنية المختارة	النافذة المثلى $\hat{h}$	$CAFPE_{opt}$	إحصائية JB	إحصائية $ARCH\text{-}LM$
NAR	2,3,7	0.4572	0.9013	34.94	44.7812
NAR-ARCH	4,5,7	0.6989	0.8671	17.78	2.2341

قمنا أولا بتحديد وتقدير نموذج الانحدار الذاتي غير الخطي بطريقة النواة، فتحصلنا على الفجوات (Lags) 2، 3 و7 وقيمة النافذة المقدرة المثلى تساوي 0.4572 عندما يكون المعيار التقاربي *CAFPE* يأخذ القيمة الصغرى 0.9013.

من الملاحظ أن توزيع بواقي نموذج NAR الممثلة في الشكل رقم (11) غير طبيعي حيث نرفض الفرضية H_0 لأن إحصائية Jarque-Bera التي تساوي 34.94 أكبر تماما من القيمة المجدولة لتوزيع χ^2 بدرجة حرية 2، فالتوزيع غير الطبيعي وغير المتناظر قد يكون دليلا على وجود بنية غير خطية، مثلا وجود أثر ARCH.

الشكل رقم (11):

بواقي تقدير النموذج NAR غير المعلمي باستعمال الفجوات 2 و 3



الجدول (14): نتائج اختبار BDS على سلسلة بواقي NAR

σ/ε		m
1	0.5	
7.5143	2.1454	2
8.1422	3.4482	3
9.8897	4.9712	4
10.5274	5.1822	5

من خلال الجدول (14)، يظهر جليا أن بواقي NAR تتميز ببنية ارتباط غير خطي لأن إحصائيات BDS أكبر تماما من القيمة الحرجة للتوزيع الطبيعي 1.96. يمكن القول إذن أن التباين الشرطي للبواقي غير متجانس باعتبار أن إحصائية ARCH-LM التي تساوي 44.7812 أكبر تماما من القيمة الحرجة لتوزيع χ^2 بدرجة حرية 1 وعليه لا بد من إضافة السيرورة ARCH إلى الانحدار الذاتي غير الخطي NAR، فيتم تقدير NAR و ARCH في آن واحد.

يُظهر الجدول (13) نتائج تحديد وتقدير نموذج الانحدار الذاتي غير الخطي مع خطأ ARCH بطريقة النواة، فتحصلنا على الفجوات (Lags) 4، 5 و7 وقيمة النافذة المقدرة المثلى تساوي 0.6989 عندما يكون المعيار التقاربي *CAFPE* يأخذ القيمة الصغرى 0.8671. نلاحظ أن إحصائية ARCH-LM على بواقي نموذج NAR-ARCH أقل تماما من القيمة الحرجة لتوزيع χ^2 بدرجة حرية 1 وهذا يعني أن التباين الشرطي لهذه البواقي متجانس، كما أنها تتميز باستقلالية تامة حيث من خلال الجدول (15)، نرفض فرضية الارتباط بين مشاهدات بواقي NAR-ARCH باعتبار أن إحصائيات BDS أقل تماما من القيمة المجدولة للتوزيع الطبيعي 1.96.

الجدول (15): نتائج اختبار BDS على سلسلة بواقي NAR-ARCH

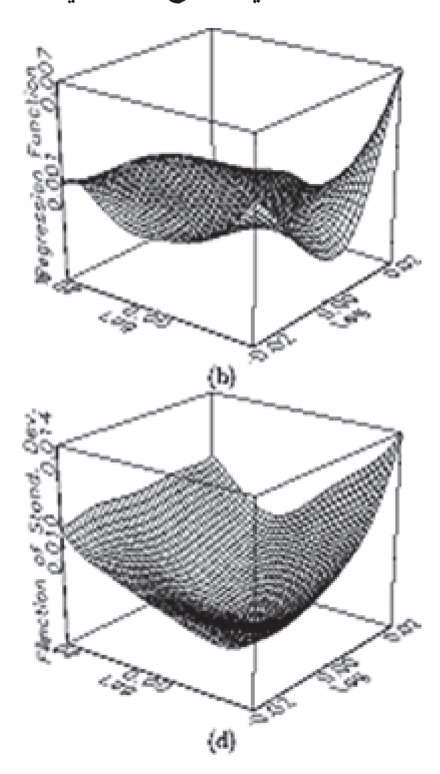
σ/ε		m
1	0.5	
1.0971	0.1434	2
1.1022	0.4182	3
1.126	0.9712	4
1.1872	1.0122	5

لتمثيل دالتي التوقع والتباين الشرطيين، نستعمل برمجية GAUSS لرسم المنحنيين على شكل 3D. ليكن البرنامج التالي:

```
Inputs:  xgrid        (xnum^2 x 2) matrix with x and y data
         f_x_grid     (xnum^2 x 2) matrix with corresponding
                       function data
         tit           character with title name
         labels       (3 x 1) character vector with axes labels
                       with y-label first, x-label second, z-label
                       last

proc(0) = plotgrid(xgrid,f_xgrid,tit,labels);
local xnum;
   title(tit);
   xlabel(labels[1]);
   ylabel(labels[2]);
   zlabel(labels[3]);
   xnum    = sqrt(rows(xgrid));
   f_xgrid = reshape(f_xgrid,xnum,xnum);
   surface(xgrid[1:xnum,2]',xgrid[1:xnum,2],f_xgrid);
endp;
```

الشكل رقم (12): تقدير دالتي التوقع الرياضي والتباين الشرطيين



من خلال الشكل رقم (12)، نلاحظ أن الدالتين هما تقريبا على شكل قطع مكافئ، هناك دليل على وجود بنية غير خطية حيث أن لسلسلة مردودية مؤشر CAC40 سلوك حركي (ديناميكي) غير متناظر أي هناك ارتباط غير خطي في التوقع الشرطي (الشكل الأعلى) والتباين الشرطي (الشكل الأسفل).

جدول التوزيع الطبيعي Laplace-Gauss

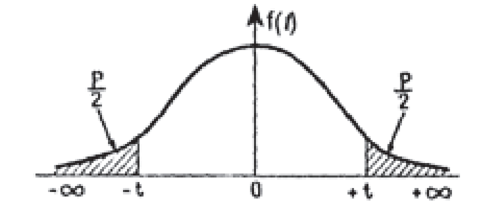


$$\pi(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}e^{-t^2/2}dt$$

t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7290	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177

t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,963?
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9779	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0.9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

جدول توزيع ستيودنت

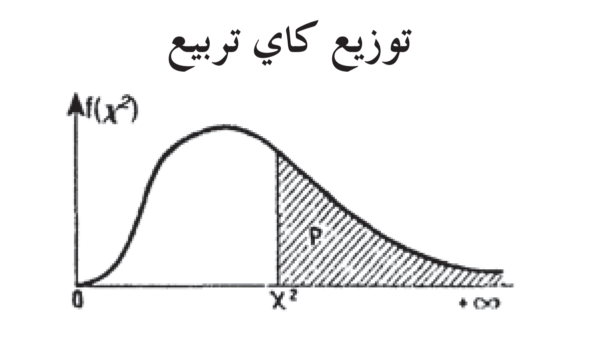


ν	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947

ν	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1.708	2,060	2,485	2,787
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750
∞	0,125	0,253	0,385	0,524	0,674	0,841	1,036	1,281	1,644	1,959	2,326	2,575

ν : عدد درجات الحرية

توزيع كاي تربيع



ν	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01
1	0,0158	0,0642	0,148	0,455	1,074	1,642	2,706	3,841	5,412	6,635
2	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	7,824	9,210
3	0,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	9,837	11,345
4	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	11,668	13,277
5	1,610	2,343	3,000	4,351	6,064	7,289	9,236	11,070	13,388	15,086
6	2,204	3,070	3,828	5,348	7,231	8,558	10,645	12,592	15,033	16,812
7	2,833	3,822	4,671	6,346	8,383	9,803	12,017	14,067	16,662	18,475
8	3,490	4,594	5,527	7,344	9,524	11,030	13,362	15,507	18,168	20,090
9	4,168	5,380	6,393	8,343	10,656	12,242	14,684	16,919	19,679	21,666
10	4,865	6,179	7,267	9,342	11,781	13,442	15,987	18,307	21,161	23,209
11	5,578	6,989	8,148	10,341	12,899	14,631	17,275	19,675	22,618	24,725
12	6,304	7,807	9,034	11,340	14,011	15,812	18,549	21,026	24,054	26,217
13	7,042	8,634	9,926	12,340	15,119	16,985	19,812	22,362	25,472	27,688
14	7,790	9,467	10,821	13,339	16,222	18,151	21,064	23,685	26,873	29,141

ν	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01
15	8,547	10,307	11,721	14,339	17,322	19,311	22,307	24,996	28,259	30,578
16	9,312	11,152	12,624	15,338	18,418	20,465	23,542	26,296	29,633	32,000
17	10,085	12,002	13,531	16,338	19,511	21,615	24,769	27,587	30,995	33,409
18	10,865	12,857	14,440	17,338	20,601	22,760	25,989	28,869	32,346	34,805
19	11,651	13,716	15,352	18,338	21,689	23.900	27,204	30,144	33,687	36,191
20	12,443	14,578	16,266	19,337	22,775	25,038	28,412	31,410	35,020	37,566
21	13,240	15,445	17,182	20,337	23,858	26,171	29,615	32,671	36,343	38,932
22	14,041	16,314	18,101	21,337	24,939	27,301	30,813	33,924	37,659	40,289
23	14,848	17,187	19,021	22,337	26,018	28,429	32,007	35,172	38,968	41,638
24	15,659	18,062	19,943	23,337	27,096	29,553	33,196	36,415	40,270	42,980
25	16,473	18,940	20,867	24,337	28,172	30,675	34,382	37,652	41,566	44,314
26	17,292	19,820	21,792	25,336	29,246	31,795	35,563	38,885	42,856	45,642
27	18,114	20,703	22,719	26,336	30,319	32,912	36,741	40,113	44,140	46,963
28	18,939	21,588	23,647	27,336	31,391	34,027	37,916	41,337	45,419	48,278
29	19,768	22,475	24,577	28,336	32,461	35,139	39,087	42,557	46,693	**49,588**
30	20,599	23,364	25,508	29,336	33,530	36,250	40,256	43,773	47,962	50,892

عندما تكون درجة الحرية ν أكبر تماما من 30، نعتبر أن العبارة $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ تخضع للقانون الطبيعي المختزل، فعلى سبيل المثال، نحسب قيمة χ^2 الموافقة للاحتمال 0.10 عندما تكون $\nu = 41$. بالاستعانة بالجدول المبين أعلاه، نحسب من أجل احتمال 0.10 و $x = 1.2816$، حيث :

$$\chi^2 = \frac{\left[x + \sqrt{2\nu - 1}\right]^2}{2} = \frac{1}{2}\left[1.2816 + \sqrt{82 - 1}\right]^2 = \frac{1}{2}(10.2816)^2 = 52.85$$

جدول توزيع فيشر

ν_2	$\nu_1 = 1$		$\nu_1 = 2$		$\nu_1 = 3$		$\nu_1 = 4$		$\nu_1 = 5$	
	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01
1	161,4	4052	199,5	4999	215,7	5403	224,6	5625	230,2	5764
2	18,51	98,49	19,00	99,00	19,16	99,17	19,25	99,25	19,30	99,30
3	10,13	34,12	9,55	30,81	9,28	29,46	9,12	28,71	9,01	28,24
4	7,71	21,20	6,94	18,00	6,59	16,69	6,39	15,98	6,26	15,52
5	6,61	16,26	5,79	13,27	5,41	12,06	5,19	11,39	5,05	10,97
6	5,99	13,74	5,14	10,91	4,76	9,78	4,53	9,15	4,39	8,75
7	5,59	12,25	4,74	9,55	4,35	8,45	4,12	7,85	3,97	7,45
8	5.32	11,26	4,46	8,65	4,07	7,59	3,84	7,01	3,69	6,63
9	5.12	10,56	4,26	8,02	3,86	6,99	3,63	6,42	3,48	6,06
10	4,96	10,04	4,10	7,56	3,71	6,55	3,48	5,99	3,33	5,64
11	4,84	9,65	3,98	7,20	3,59	6,22	3,36	5,67	3,20	5,32
12	4.75	9,33	3,88	6,93	3,49	5,95	3,26	5,41	3,11	5,06
13	4.67	9,07	3,80	6,70	3,41	5,74	3,18	5,20	3,02	4,86
14	4.60	8,86	3,74	6,51	3,34	5,56	3,11	5,03	2,96	4,69

ν_2	$\nu_1 = 1$		$\nu_1 = 2$		$\nu_1 = 3$		$\nu_1 = 4$		$\nu_1 = 5$	
	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01
15	4,54	8,68	3,68	6,36	3,29	5,42	3,06	4,89	2,90	4,56
16	4.49	8,53	3,63	6,23	3,24	5,29	3,01	4,77	2,85	4,44
17	4.45	8,40	3,59	6,11	3,20	5,18	2,96	4,67	2,81	4,34
18	4.41	8,28	3,55	6,01	3,16	5,09	2,93	4,58	2,77	4,25
19	4.38	8,18	3,52	5,93	3,13	5,01	2,90	4,50	2,74	4,17
20	435	8,10	3,49	5,85	3,10	4,94	2,87	4,43	2,71	4,10
21	4.32	8,02	3,47	5,78	3,07	4,87	2,84	4,37	2,68	4,04
22	4.30	7,94	3,44	5,72	3,05	4,82	2,82	4,31	2,66	3,99
23	4.28	7,88	3,42	5,66	3,03	4,76	2,80	4,26	2,64	3,94
24	4,26	7,82	3,40	5,61	3,01	4,72	2,78	4,22	2,62	3,90
25	4,24	7,77	3,38	5,57	2,99	4,68	2,76	4,18	2,60	3,86
26	4,22	7,72	3,37	5,53	2,98	4,64	2,74	4,14	2,59	3,82
27	4.21	7,68	3,35	5,49	2,96	4,60	2,73	4,11	2,57	3,78
28	4,20	7,64	3,34	5,45	2,95	4,57	2,71	4,07	2,56	3,75
29	4.18	7,60	3,33	5,42	2,93	4,54	2,70	4,04	2,54	3,73
30	4.17	7,56	3,32	5,39	2,92	4,51	2,69	4,02	2,53	3,70
40	4.08	7,31	3,23	5,18	2,84	4,31	2,61	3,83	2,45	3,51
60	4.00	7,08	3,15	4,98	2,76	4,13	2,52	3,65	2,37	3,34
120	3,92	6,85	3,07	4,79	2,68	3,95	2,45	3,48	2,29	3,17
∞	3.84	6,64	2,99	4,60	2,60	3,78	2,37	3,32	2,21	3,02

ν_1 : عدد درجات الحرية الخاصة بالبسط

ν_2 : عدد درجات الحرية الخاصة بالمقام

جدول توزيع فيشر

ν_2	$\nu_1 = 6$		$\nu_1 = 8$		$\nu_1 = 12$		$\nu_1 = 24$		$\nu_1 = \infty$	
	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01
1	161,4	4052	199,5	4999	215,7	5403	224,6	5625	230,2	5764
2	18,51	98,49	19,00	99,00	19,16	99,17	19,25	99,25	19,30	99,30
3	10,13	34,12	9,55	30,81	9,28	29,46	9,12	28,71	9,01	28,24
4	7,71	21,20	6,94	18,00	6,59	16,69	6,39	15,98	6,26	15,52
5	6,61	16,26	5,79	13,27	5,41	12,06	5,19	11,39	5,05	10,97
6	5,99	13,74	5,14	10,91	4,76	9,78	4,53	9,15	4,39	8,75
7	5,59	12,25	4,74	9,55	4,35	8,45	4,12	7,85	3,97	7,45
8	5.32	11,26	4,46	8,65	4,07	7,59	3,84	7,01	3,69	6,63
9	5.12	10,56	4,26	8,02	3,86	6,99	3,63	6,42	3,48	6,06
10	4,96	10,04	4,10	7,56	3,71	6,55	3,48	5,99	3,33	5,64
11	4,84	9,65	3,98	7,20	3,59	6,22	3,36	5,67	3,20	5,32
12	4.75	9,33	3,88	6,93	3,49	5,95	3,26	5,41	3,11	5,06
13	4.67	9,07	3,80	6,70	3,41	5,74	3,18	5,20	3,02	4,86
14	4.60	8,86	3,74	6,51	3,34	5,56	3,11	5,03	2,96	4,69

ν_2	$\nu_1 = 6$		$\nu_1 = 8$		$\nu_1 = 12$		$\nu_1 = 24$		$\nu_1 = \infty$	
	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01
15	4,54	8,68	3,68	6,36	3,29	5,42	3,06	4,89	2,90	4,56
16	4.49	8,53	3,63	6,23	3,24	5,29	3,01	4,77	2,85	4,44
17	4.45	8,40	3,59	6,11	3,20	5,18	2,96	4,67	2,81	4,34
18	4.41	8,28	3,55	6,01	3,16	5,09	2,93	4,58	2,77	4,25
19	4.38	8,18	3,52	5,93	3,13	5,01	2,90	4,50	2,74	4,17
20	435	8,10	3,49	5,85	3,10	4,94	2,87	4,43	2,71	4,10
21	4.32	8,02	3,47	5,78	3,07	4,87	2,84	4,37	2,68	4,04
22	4.30	7,94	3,44	5,72	3,05	4,82	2,82	4,31	2,66	3,99
23	4.28	7,88	3,42	5,66	3,03	4,76	2,80	4,26	2,64	3,94
24	4,26	7,82	3,40	5,61	3,01	4,72	2,78	4,22	2,62	3,90
25	4,24	7,77	3,38	5,57	2,99	4,68	2,76	4,18	2,60	3,86
26	4,22	7,72	3,37	5,53	2,98	4,64	2,74	4,14	2,59	3,82
27	4.21	7,68	3,35	5,49	2,96	4,60	2,73	4,11	2,57	3,78
28	4,20	7,64	3,34	5,45	2,95	4,57	2,71	4,07	2,56	3,75
29	4.18	7,60	3,33	5,42	2,93	4,54	2,70	4,04	2,54	3,73
30	4.17	7,56	3,32	5,39	2,92	4,51	2,69	4,02	2,53	3,70
40	4.08	7,31	3,23	5,18	2,84	4,31	2,61	3,83	2,45	3,51
60	4.00	7,08	3,15	4,98	2,76	4,13	2,52	3,65	2,37	3,34
120	3,92	6,85	3,07	4,79	2,68	3,95	2,45	3,48	2,29	3,17
∞	3.84	6,64	2,99	4,60	2,60	3,78	2,37	3,32	2,21	3,02

ν_1 : عدد درجات الحرية الخاصة بالبسط

ν_2 : عدد درجات الحرية الخاصة بالمقام

جدول دربين–واتسون

n	$k = 1$		$k = 2$		$k = 3$		$k = 4$		$k = 5$	
	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,10	1,37	0,98	1,54	0,82	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,90	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74	1,07	1,83
31	1,36	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82
33	1,38	1,51	1,32	1,58	1,26	1,65	1,19	1,73	1,13	1,81

n	$k=1$		$k=2$		$k=3$		$k=4$		$k=5$	
	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2
34	1,39	1,51	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80
36	1,41	1,52	1,35	1,59	1,29	1,65	1,24	1,73	1,18	1,80
37	1,42	1,53	1,36	1,59	1,31	1,66	1,25	1,72	1,19	1,80
38	1,43	1,54	1,37	1,59	1,32	1,66	1,26	1,72	1,21	1,79
39	1,43	1,54	1,38	1,60	1,33	1,66	1,27	1,72	1,22	1,79
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77
55	1,53	1,60	1,49	1,64	1,45	1,68	1,41	1,72	1,38	1,77
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77
65	1,57	1,63	1,54	1,66	1,50	1,70	1,47	1,73	1,44	1,77
70	1,58	1,64	1,55	1,67	1,52	1,70	1,49	1,74	1,46	1,77
75	1,60	1,65	1,57	1,68	1,54	1,71	1,51	1,74	1,74	1,77
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77
85	1,62	1,67	1,60	1,70	1,57	1,72	1,55	1,75	1,52	1,77
90	1,63	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78
95	1,64	1,69	1,62	1,71	1,60	1,73	1,58	1,75	1,56	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

n : حجم العينة (عدد المشاهدات)

k : عدد المتغيرات المستقلة في النموذج (الثابتة مقصاة)

جداول ديكي- فولر

جداول توزيع $t_{\hat{\phi}_1}$

نوع النموذج	عدد المشاهدات n	الاحتمالات							
		0,01	0,025	0,05	0,10	0,90	0,95	0,975	0,99
النموذج 1	25	-2,66	-2,26	-1,95	-1,60	0,92	1,33	1,70	2,16
	50	-2,62	-2,25	-1,95	-1,61	0,91	1,31	1,66	2,08
	100	-2,60	-2,4	-1,95	-1,61	0,91	1,29	1,64	2,03
	250	-2,58	-2,23	-1,95	-1,62	0,89	1,29	1,63	2,01
	500	-2,58	-2,23	-1,95	-1,62	0,89	1,28	1,62	2,00
	∞	-2,58	-2,23	-1,95	-1,62	0,89	1,28	1,62	2,00
النموذج 2	25	-3,75	-3,33	-3,00	- 2,63	-0,37	0,00	0,34	0,72
	50	-3,58	-3,22	-2,93	-2,60	-0,40	-0,03	0,29	0,66
	100	-3,51	-3,17	-2,89	-2,58	-0,42	-0,05	0,26	0,63
	250	-3,46	-3,14	-2,88	-2,57	-0,42	-0,06	0,24	0,62
	500	-3,44	-3,13	-2,87	-2,57	-0,43	-0,07	0,24	0,61
	∞	-3,43	-3,12	-2,86	-2,57	-0,44	-0,07	0,23	0,60
النموذج 3	25	-4,38	-3,95	-3,60	-3,24	-1,14	-0,80	-0,50	-0,15
	50	-4,15	-3,80	-3,50	-3,18	-1,19	-0,87	-0,58	-0,24
	100	-4,04	-3,73	-3,45	-3,15	-1,22	-0,90	-0,62	-0,28

نوع النموذج	عدد المشاهدات n	الاحتمالات							
		0,01	0,025	0,05	0,10	0,90	0,95	0,975	0,99
	250	-3,99	-3,69	-3,43	-3,13	-1,23	-0,92	-0,64	-0,31
	500	-3,98	-3,68	-3,42	-3,13	-1,24	-0,93	-0,65	-0,32
	∞	-3,96	-3,66	-3,41	-3,12	-1,25	-0,94	-0,66	-0,33

النموذج 1 :بدون اتجاه عام وبدون حد ثابت

النموذج 2 : بدون اتجاه عام ولكن يتضمن حدا ثابتا

النموذج 3 : يتضمن اتجاها عاما وحدا ثابتا

جداول توزيع $t_{\hat{c}}$ و $t_{\hat{b}}$

n	النموذج 2			النموذج 3					
	الثابتة c			الثابتة c			الاتجاه b		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
100	3,22	2,54	2,17	3,78	3,11	2,73	3,53	2,79	2,38
250	3,19	2,53	2,16	3,74	3,09	2,73	3,49	2,79	2,38
500	3,18	2,52	2,16	3,72	3,08	2,72	3,48	2,78	2,38
∞	3,18	2,52	2,16	3,71	3,08	2,72	3,46	2,78	2,38

المراجـــع

باللغة العربية:

1- امتثال محمد حسن، محمد علي محمد أحمد، مبادئ الاستدلال الإحصائي، الإسكندرية: الدار الجامعية، 2000.

2- السعيد هتهات، دراسة اقتصادية وقياسية لظاهرة التضخم في الجزائر، مذكرة لنيل شهادة الماجستير، غير منشورة، جامعة ورقلة، كلية الحقوق والعلوم الاقتصادية، 2006.

3- المرسي السيد الحجازي، عبد القادر محمد عطية، مقدمة في الاقتصاد القياسي: المبادئ والتطبيقات، الرياض: النشر العلمي والمطابع، 2001.

4- أموري هادي كاظم الحسناوي، طرق القياس الاقتصادي، عمان: دار وائل للنشر، 2002.

5- أيت طالب حميد، محاولة بناء نموذج اقتصادي للتضخم في الجزائر، مذكرة لنيل شهادة الماجستير، غير منشورة، جامعة الجزائر، كلية العلوم الاقتصادية، 1997.

6- تومي ربيعة، نمذجة سعر الصرف الاسمي في المدى الطويل باستعمال طريقة التكامل المشترك، مذكرة لنيل شهادة الماجستير، غير منشورة، جامعة الجزائر، كلية العلوم الاقتصادية وعلوم التسيير، 2002.

7- تومي صالح ، النمذجة القياسية للتضخم في الجزائر خلال الفترة 1988–2000، أطروحة دكتوراه الدولة، غير منشورة، جامعة الجزائر، كلية العلوم الاقتصادية وعلوم التسيير، 2002.

8- تومي صالح، مدخل لنظرية القياس الاقتصادي، الجزائر: ديوان المطبوعات الجامعية، ج(2)، 1999.

9- تومي صالح، مدخل لنظرية القياس الاقتصادي، الجزائر: ديوان المطبوعات الجامعية، ج(1)، 1999.
10- حسن محمد حسن محمد، أساسيات الإحصاء وتطبيقاته، الإسكندرية: دار المعرفة الجامعية، بدون سنة.
11- حشمان مولود، نماذج وتقنيات التنبؤ القصير المدى، الجزائر: ديوان المطبوعات الجامعية، 2002.
12- حشمان مولود، محددات الأجر في الجزائر، أطروحة دكتوراه الدولة، غير منشورة، جامعة الجزائر، كلية العلوم الاقتصادية وعلوم التسيير، 2000.
13- زياد رمضان، مبادئ الإحصاء الوصفي والتطبيقي و الحيوي، طب 5؛ عمان: دار وائل للنشر، 2001.
14- السعيد بومنجل، الدليل الإحصائي للطالب، الجزائر: ديوان المطبوعات الجامعية، 2000.
15- سلفادور دومينيك، الإحصاء والاقتصاد القياسي، طب2 ؛ الجزائر: ديوان المطبوعات الجامعية، 1993.
16- سمير محمد عبد العزيز، الاقتصاد القياسي: مدخل في اتخاذ القرارات، الإسكندرية: مكتبة الإشعاع للطباعة والنشر والتوزيع، 1997.
17- شرابي عبد العزيز، طرق إحصائية للتوقع الاقتصادي، الجزائر: ديوان المطبوعات الجامعية، 2000.
18- صالح تركي القريشي، ناظم محمد نوري الشمري، مبادئ علم لاقتصاد، الموصل: دار الكتب للطباعة والنشر، 1993.
19- عبد الحميد عبد المجيد البلداوي، الإحصاء للعلوم الإدارية والتطبيقية، عمان: دار الشروق للنشر والتوزيع، 1997.

20- عبد الرحمان بن محمد سليمان أبو عمه، أنور أحمد محمد عبد الله، محمود محمد إبراهيم هنيدي، الإحصاء التطبيقي، الرياض: مطابع جامعة الملك سعود، 1995.

21- عبد القادر محمد عبد القادر عطية، الاقتصاد القياسي بين النظرية والتطبيق، ط(2) ؛ الإسكندرية: الدار الجامعية، 2000.

22- عبد القادر محمد عبد القادر، طرق قياس العلاقات الاقتصادية مع تطبيقات على الحاسب الالكتروني، الإسكندرية: دار الجامعات المصرية، 1990.

23- عبد القادر محمد عبد القادر، طرق قياس العلاقات الاقتصادية مع تطبيقات على الحاسب الالكتروني، الإسكندرية: دار الجامعات المصرية، 1990.

24- عبد المنعم السيد علي، نزار سعد الدين العيسي، النقود والمصارف والأسواق المالية، ط(1) ؛ عمان: دار الحامد للنشر والتوزيع، 2004.

25- عبد الناصر العبادي، عبد الحليم كراجة، محمد الباشا، مبادئ الاقتصاد الكلي، ط(1) ؛ عمان: دار صفاء للنشر والتوزيع، 2000.

26- قبلي زهير، تحديد سعر النفط الخام في الأجلين القصير والطويل باستعمال تقنيات التكامل المتزامن ونماذج تصحيح الخطأ، مذكرة لنيل شهادة الماجستير، غير منشورة، جامعة الجزائر، كلية العلوم الاقتصادية وعلوم التسيير، 1999.

27- كمال سلطان محمد سالم، الإحصاء الاحتمالي، الإبراهيمية: الدار الجامعية، 2004.

28- لبيبة حسب النبي العطار، مقدمة في الاستدلال الإحصائي، الإسكندرية: الدار الجامعية للطباعة والنشر والتوزيع، 1993.

29- لزعر علي، الإحصاء وتوفيق المنحنيات، الجزائر: ديوان المطبوعات الجزائرية، 2000.

30- مجدي محمود شهاب، الاقتصاد النقدي، الإسكندرية: الدار الجامعية ،1990.

31- مصطفى الخواجة، مقدمة في الإحصاء، الإسكندرية: الدار الجامعية، 2002.

32- ناظم حيدر، الوسيط في الإحصاء التطبيقي، طب 2؛ دمشق: دار الكتاب، 1977.

33- نصيب رجم، الإحصاء التطبيقي، عنابه: دار العلوم للنشر والتوزيع، 2004.

34- نعمة الله نجيب إبراهيم، مقدمة في مبادئ الاقتصاد القياسي، الإسكندرية: مؤسسة شباب الجامعة، 2002.

35- هني أحمد، دروس في التحليل الاقتصادي الكلي، الجزائر: ديوان المطبوعات الجامعية،1991.

36- وليد اسماعيل السيفو و أحمد محمد مشعل، الاقتصاد القياسي التحليلي بين النظرية و التطبيق، عمان: دار مجدلاوي للنشر و التوزيع، 2003.

باللغات الأجنبية:

37- Abraham-Frois G. (1994), « La dynamique chaotique », Sirey.

38- Akaike, H. (1973), « Information theory and an extension of the maximum likelihood prin-ciple », in B.N. Petrov and F. Csâki, eds, *2nd International Symposium on Information Theory,* Akadémia Kiado, Budapest.

39- Akaike, H. (1974), « A new look at the statistical model identification », IEEE *Transactions on automatic Control,* Vol. 19.

40- Akaike, H. (1979), « A bayesian extension of the minimum AIC procédure », *Biometrika,* Vol. 66.

41- Alexander, J. (1961), “Price movements in speculative markets; trends or random walk”.Industrial Management Review.
42- Alexandre, H et Ertur, K.C. (1994), « Impact de l'intervalle d'échantillonnage sur les tests d'efficience: application au marché français des actions », Finance, 15, 7-27.
43- Alexandre, H. (1992), « La quasi marche aléatoire », Finance, 13,2, 5-21
44- ALMON S. (1965), « The distributed lag between capital appropriation and expenditures », *Econometrica,* Vol. 33, n° l.
45- Altman, N.S. (1990), “Kernel smoothing of data with correlated errors”. J. Am. Statist. Assoc, 85, 749-759.
46- Andersen, T. G., T, Bollerslev, P. F., Christo®ersen and Diebold, F.X. (2006), “Volatility and correlation forecasting”. In G. Elliott, C. W. J. Granger, and A. Timmermann (eds.), Handbook of Economic Forecasting, Amsterdam: North-Holland, 778{878).
47- Andrews, D. (1991), “Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation”, Econometrica, 59, 817-858.
48- Andrews, D. W. K., and Guggenberger, P. (2003), “A Bias-Reduced Log-Periodogram Regression Estimator for the Long-memory Parameter”, Econometrica, 71, 675-712.
49- Auestad, B. and Tjostheim, D. (1990), “Identification of nonlinear time series: first order characterization and order determination”, Biometrika, 77,4, 669-687.
50- Baillie, R.T. (1996), “Long memory processes and fractional integration in Econometrics”, Journal of Econometrics, 73, 5-59
51- Baillie, R.T and Chung, C.F. (1996), “Analysing inflation by the fractionally integrated ARFIMA-GARCH model”, Journal of Econometrics, 11, 23-40.
52- Baillie, R.T., Bollerslev, T, and Mikkelsen, H.O. (1996), “Fractionally integrated generalized autoregressive conditional heteroskedasticity”. J. of Econometrics, 74, 3-30.
53- Baillie, R.T., Chung, C.F. and Tiles, M.A. (1995), “Analyzing inflation by the fractionally integrated ARFIMA-GARCH model”. J. of Appl. Econometrics, 11, 23–40.
54- Ball, R and Brown, P. (1968), “An empirical evaluation of accounting income numbers”. Journal of Accounting Research.
55- Bartlett, M.S, (1990),:“Chance or chaos?”, J. R. Statist. Soc. B, Vol. 153, Part 3, pp. 321-347.
56- Bendib, R. (2001), « Econométrie: Théorie et Applications ». Alger, O.P.U..

57- Beran, J. (1994), "Statistics for long-memory processes". Chapman and Hall, New York.
58- Beran, J. (1995), "Maximum likelihood estimation of the differencing parameter for invertible short- and long-memory ARIMA models". J. Roy. Statist. Soc. B, 57, 672-695.
59- Beran, J., Bhansali, R.J., Ocker, D. (1998), "On unified model selection for stationary and nonstationary short-and long-memory autoregressive processes". Biometrika, 85, 921–934.
60- Beran, J. and Feng, Y. (2001), "Local polynomial estimation with a FARIMA-GARCH error process". Bernoulli, 7, 733–750.
61- Beran, J. and Feng, Y. (2002a), "SEMIFAR models - A semiparametric framework for modelling trends, long-range dependence and nonstationarity". Computat. Statist. Data Anal., 40, 393–419.
62- Reveveridge, S and Nelson, C.R. (1981), "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle", Journal ofthe Monetary Economies, 1, 2, 1981.
63- Bierens, H.J. (1987), "Kernel estimators of regression functions". In Advances in Econometrics, 99-144, Cambridge University Press, Cambridge.
64- Blomquist, S, Eklof, M and Newey, W. (1997), "Tax reform using nonparametric methods". Sweden 1980-1991. Uppsala- Working papers series.
65- Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", Journal of Econometrics, 31, 307-327.
66- Bollerslev, T, Chou, R.Y, Jayaraman N. et Kroner K.F. (1991), « Les modèles ARCH en finance: un point sur la théorie et les résultats empiriques », Annales d'économie et de statistique, 24,1-59.
67- Booth, G. G, Kaen, F.R et Koveos, P.E. (1982), "R/S Analysis of Foreign Exchange Rates under Two International Monetary Regimes", Journal of Monetary Economics, 407-415.
68- Bosq, D et Lecoutre, J.P. (1987), « Théorie de l'estimation fonctionnelle ». Economica, Paris.
69- Bosq, D et Lecoutre, J.P. (1992), « Analyse et prévision des séries chronologiques ». Masson, Paris
70- Bosq, D. (1979), « Sur la prédiction non paramétrique de variables aléatoires et mesures aléatoires », Pub. Interne, UER de Mathématiques, Lilies.

71- Bosq, D. (1996), “Nonparametric statistics for stochastic processes”. Lecture Notes in statistics, 110, Springer-verlag.
72- Bourbonnais, R. (2003), « Econométrie ». 5e édition. Paris, Dunod, 2003.
73- Bourbonnais, R et Terraza, M. (1998), « Analyse des séries temporelles en économie ». Paris, PUF.
74- Box G.E.P., Jenkins G.M. (1976), “Time series analysis: forecasting and control”, Holdenday.
75- Box, G.E.P., Pierce D.A. (1970), « Distribution of residual autocorrelations in autoregressive-integrated moving average time series models », Journal of the American Statistical Association, Vol. 65.
76- Bresson, G et Michaud, G.C. (1995), « Econométrie des séries temporelles Théorie et application ». Paris, P.U.F.
77- Breusch, T. (1978), « Testing for autocorrelation in dynamic linear models », Australian Economie Papers, Vol. 17.
78- Brock, W.A and Baek, E.G. (1991), “Some Theory of Statistical Inference for Nonlinear Science”, Review of Economic Studies, 58, 697-716.
79- Brock, W.A and Hommes, C.H. (1998), “Heterogeneous beliefs and routes to chaos in a simple asset pricing model”, Journal of Economic Dynamics & Control, Vol. 22, pp. 1235-1274.
80- Brock, W.A, Dechert, W.D. et Scheinkman, J.A. (1987), “A Test for Independence Based on the Correlation Dimension”, Working Paper, University of Wisconsin.
81- Brock, W.A, Hsieh, D.A and LeBaron, B. (1992), “Nonlinear Dynamics, Chaos and Instability”, MIT Press, 328 pages, second edition.
82- Brock, W.A, Dechert, W.D, Scheinkman, J.A and LeBaron, B. (1996),
83- Brockmann, M. (1993), “Locally adaptive bandwidth choice for kernel regression estimators”. J. Amer. Statist. Assoc, 88,1302-1309.
84- Brockwell, P.J and Davis, R. (1996), “Introduction to time series and forecasting”, Springer-Verlag, 1996.
85- Brown, R.G, Durbin, J and Evans, J.M. (1975), « Techniques for testing the constancy of the regression relationship overtime », Journal of the Royal Statistical Society, B, 37(2).
86- Camlong, C, Sarda, P and Vieu, P. (1998), “Additive time series: The kernel integration method”, Univ-Paul Sabatier Toulouse I.

87- Campbell, J.Y and Shiller, R.J. (1987), “Cointegration and Tests for Present Value Models”. Journal of Political Economy, 95(5), 1062-1088.

88- Chan, K.S. and Tong, H., (1994): “A note on noisy chaos”, J. R. Statist. Soc. B, Vol. 2, pp. 301-311.

89- Chen, R, Hardie, W and Linton, O and Severance-Lossin, E. (1995), “Nonparametric estimation of additive separable regression models”, Discussion paper 95-50, SFB 373, Humboldt Universitât zu Berlin.

90- Chen, S.H., Lux, T. and Marchese, M., (1999): “Testing for non-linear structure in an artificial financial market”, discussion paper B-447, University of Bonn.

91- Cheng, B. and Tong, H. (1992), “On consistent nonparametric order determination and chaos”, Journal of The Royal Statistical Society, Series B, 54,427-449.

92- Cheung, Y.W. (1993a), “Tests for Fractional Integration: A Monte Carlo Investigation”. Journal of Time series Analysis, 14(4), 331-345.

93- Cheung, Y.W. (1993b), “Long Memory in Foreign Exchange Rates”. Journal of Business and Economic Statistics, 11(1), 93-101.

94- Chikhi, M and Diebolt, C. (2009), "The Reichsbank: A Nonparametric Modelling of Historical Time Series", in: Applied Financial Economics Letters (Routledge)

95- Chikhi, M et Terraza, M. (2001), « Prévision non paramétrique de l'action France Télécom ». Working paper, LAMETA.

96- Chikhi, M. (2001), « Le marché boursier en France est-il efficient ? Application à la prévision non paramétrique de l'indice CAC40 ». Working paper, LAMETA.

97- Chikhi, M, (2001), "Modélisation non paramétrique des processus stochastiques: Analyse non paramétrique de non linéarité de l'indice CAC40 ". Thèse de doctorat, Université de Montpellier I.

98- Chili, S.T. (1989), “Bandwidth selection for kernel estimates with correlated noise”. Statist. Probab. Lett., 8, 347-354.

99- Gourieroux, C. (1992), « Modèles ARCH et application financière, Paris: Economica.

100- Chu, C. K and Marron, S. (1991), “Comparison of two bandwidth selectors with dependent errors”. Annal. Statist. 19,1906-1918.

101- Chung, CF. (1994), “A Note on Calculating the Autocovariances of the Fractionally Integrated ARMA Models”, Economics Letters, 45(3), 293-297.

102- Cochrane, J.H. (1988), “How Big is the Random Walk in GNP ?”, Journal of Political Economy, 96, 893-920.
103- Cochrane, J.H. (1991), “Volatility Tests and Efficient Markets: A Review Essay”, Journal of Monetary Economics, 27,463-485.
104- Collomb, G. (1976), « Estimation non paramétrique de la régression par la méthode du noyau », Thèse de l'université de Toulouse I.
105- Collomb, G and Hardie, W. (1986), “Strong mixing convergence rates in robust nonparametric time series and prediction for dependent observations”, Stoch. Proc. and their Appl. 23, 77-89.
106- Collomb, G. (1979), « Conditions nécessaires et suffisantes de convergence uniforme d'un estimateur de la régression, estimation des dérivées de la régression. C.R.A.S., Ser.. A., 288, 161-163.
107- Collomb, G. (1980), Estimation non paramétrique de probabilités conditionnelles, C.R. Acad. sci. Paris Sér I Math. 291,427-430.
108- Collomb, G. (1981), « Prédiction non paramétrique: étude de l'erreur quadratique du prédictogramme », Pub. Interne, Univ. P. Sabatier, Toulouse.
109- Collomb, G. (1984), « Propriétés de convergence presque complète du prédicteur à noyau », Zeitschrift fur Wahschein lichkeitstheorie und verwandte Gedbiete, 66,441-460.
110- Collomb, G. (1985,), “Nonparametric regression”. An up to date bibliography. Statistics, 16(2), 309-324.
111- Crato, N and Rothman, P. (1994), “Fractional Integration Analysis of Long-Run Behavior for US Macroeconomics Time Series”, Economics Letters, 45(3), 287-291.
112- Csôrgô, S.and Mielniczuk, J. (1995), “Nonparametric regression under long-range dependent normal errors. Annals of Statistics, 23, 1000-1014.
113- David et Michaud, J. (1989), « La Prévision: Approche empirique d'une méthode statistique, Paris: Masson.
114- Davidson, J. (2001), “Moment and Memory Proprieties of Linear Conditional Heteroscedasticity Models”. Manuscript, Cardiff University.
115- Davidson, J. (2004), “Moment and memory properties of linear conditional heteroscedasticity models, and a new model”. Journal of Business & Economic Statistics 22, 16{19.
116- Davidson, J., Terasvirta, T.T. (Eds.), (2002), “Long Memory and Nonlinear Time Series”, Journal of Econometrics, 110 (2) 105–437.
117- Dechert, W.D and Gençay, R., (1990), “Estimating Lyapunov exponents with mulrilayer feedforward network learning”, Working paper, Department of Economics, University of Houston.

118- Deheuvels, P. (1974), « Conditions nécessaires et suffisantes de convergence presque sûre et uniforme presque sûre des estimateurs de la densité ». C.R. Acad. Sci. Paris A, 278, 1217rl220.
119- Deheuvels, P. (1977), « Estimation non paramétrique de la densité par histogramme généralisé », Revue de Statistique Appliquée, 35, 5-42.
120- Devroye, L. (1983), "The equivalence of weak, strong and complete convergence in I) for kernel density estimates", Ann. Stetist. 12, 1231-1249.
121- Dickey, D and Fuller, W. (1979), "Distribution of the estimators for autoregressive time séries with unit root". Journal of the American Statistical Association, Vol. 74, n°366.
122- Dickey, D and Fuller, W. (1981), "Likelihood ratio statistics for autoregressive time séries with unit root". Econometrica, Vol. 49, n°4, 1981.
123- Doornik, J.A., Ooms, M., (2004), "Inference and forecasting for ARFIMA models with an application to US and UK inflation". Studies in Nonlinear Dynamics and Econometrics, 8, No. 2, Article 14.
124- Dormont, B. (1999), « Introduction à l'économétrie », Montchrestien, Paris, 1999.
125- Doukhan, P. (1994,), "Mixing: Properties and examples". New-York; Springer-Verlag.
126- Droesbeke, J.J, Fichet, B and Tassi, P. (1994), « Modélisation ARCH: Théorie statistique et applications dans le domaine de la finance ». Belgique: Editions de L'universite de Bruxelles.
127- Durbin, J. (1970), "Testing for serial correlation in least-squares regression when some of the regressors are lagged dependant variables", Biometrika, Vol. 38.
128- Durbin, J and Watson, G.S. (1951), "Testing for serial correlation in least-squares regression", Biometrika, Vol. 38.
129- Engle, R.E and Granger, C.W.J. (1987), "Cointegration and Error-correction: representation, estimation and testing", Econometrica, Vol. 55.
130- Engle, R.E, Hendry, D.F and Richard, J.F. (1983), "Exogeneity". Econometrica, Vol. 51.
131- Engle, R.F. (1982), "Autoregressive conditional heteroscedasticity with estimate of the variance of U.K. inflation ». Econometrica, Vol. 50.

132- Engle, R.F., Lilien, D.M. and Robins, R.P. (1987), "Estimating time-varying risk premia in the term structure: the ARCH-M model". Econometrica, 55, 391–407.
133- Farmer, D.J. and Sidorowitch, J.J. (1987), "Predicting chaotic time series", Physical Review Letters 59, pp. 845-848.
134- Farrar, D.E and Glauber, R.R. (1967), "Multicolinearity in regression analysis". Review of Economies and Statistics, Vol. 49.
135- Feng, Y. (2004), "Non- and Semiparametric Regression with Fractional Time Series Errors – Theory and Applications to Financial Data". Habilitation Monograph, University of Konstanz.
136- Fomby, T.B and Carter Hill, R and Johnston, S.R. (1984), "Advanced econometric methods", Springer-Verlag.
137- Fox, R and Taqqu, M.S. (1986), "Large Sample properties of Parameter Estimates for Strongly Dependant Stationary Gaussian Time Series". Annals of Statistics, 14, 517.
138- Fuller, W.A. (1976), "Introduction to statistical time series". John Wiley.
139- Gençay, R. and Dechert, W.D. (1992), "An algorithm for the Lyapunov exponents of an n-dimensional unknown dynamical system", Physica D, Vol. 59, pp. 142-157.
140- Gençay, R and Liu, T. (1996), "Nonlinear modelling and prediction with feedforward and recurrent networks". Working paper, Department of Economics, University of Windsor, Canada.
141- Geweke, J and Porter-Hudak, S. (1983), "The Estimation and Application of Long Memory Time Series Models". Journal of Time Series Analysis, 4(4), 221-238.
142- Gillet, P. (1999), « L'efficience des marchés financiers ». Economica.
143- Gleisier, H. (1969), "A new test for heteroscedasticity". Journal of American Statistical Association,Vol. 64.
144- Godfrey, L.G. (1978), "Testing for higher order serial correlation in regression equation when the regressors contain lagged dependant variables", Econometrica, Vol. 46.
145- Goldfeld, S.M and Quandt, R.E. (1965), "Some tests of homoscedasticity". Journal ofthe American Statistical Association, Vol. 60.
146- Goldfeld, S.M and Quandt, R.E. (1972), "Non linear methods in econometrics". North-Holland. Amsterdam.
147- Goodwin, R.M. (1951), The Non-Linear Accelerator and the Persistence of Business Cycles, Econometrica, 19,1-17.

148- Gourierous, C et Manfort, A. (1983), « Cours de séries temporelles ». Paris. Economica.
149- Gourieroux, C et Monfort, A. (1995), « Séries temporelles et modèles dynamiques, Paris: Economica.
150- Grais, B. (1978), « Méthodes Statistiques ». Paris, Dunod.
151- Granger, C.W.J. (1983), "Co-integrated variables and Error-correcting models". Working paper. University of San Diego.
152- Granger, C.W.J. (1969), "Investigating causal relations by econometrics models and cross spectral methods". Econometrica, Vol. 37.
153- Granger, C.W.J and Joyeux, R. (1980), "An introduction to long memory time series and fractional differencing", Journal of Time Series Analysis, 1,1-15.
154- Grassberger, P and Procaccia, I., (1983a), "Measuring the strangeness of strange attractors", Physica 9D, pp. 189-208.
155- Greene, W.H. (2000), "Econometric Analysis", Prentice Hall, 4e ed.
156- Griffiths, W.E and Carter Hill, R and Judge, G.G. (1993), "Learning and practicing econometrics". John Willey, New York.
157- Guégan, D. (1994), « Séries chronologiques non linéaire à temps discret », Economica.
158- Hafner, C. (1996), "Estimating High Frequency Foreign Exchange Rate Volatility with Nonparametric ARCH Models". Discussion paper 96-17, SFB 373, Humboldt Universitât zu Berlin.
159- Hall, P and Hart, J.D. (1990), "Nonparametric regression with long-range dependence". Stochastic Processes and their application, 36,339-351
160- Hamdani, H. (1988), « Statistique descriptive et expression graphique ». Alger, OPU.
161- Hamid, K, Khenouse, A et Zatout, A. (1998), «Modèles Autorégrssifs Conditionnellement Hétéroscédastique ». Revue d'Economie et de Statistique Appliquée, INPS, N°0, Alger.
162- Hamilton, J.D. (1994), "Time series analysis". Princeton University Press.
163- Hannan, E.J. (1973), "The asymptotic theory of linear time series models". J. of Appl. Prob. , 10, 130-145.
164- Hannan, E.J and Quinn, B.G. (1979), "The determination of the order of an autoregression". Journal of the royal statistical Society Series B, 41,190-195.
165- Hardle, W and Chen, R. (1996), "Nonparametric Time Series Analysis, a selective review with examples". Proceedings of the 50th session of the ISI, Peking.

166- Hardle, W and Linton, O. (1994), “Applied nonparametric methods”. The Handbook of Econometrics, vol IV, Elsevier Sience. B.V.
167- Hardle, W and Linton, O. (1993), “Applied nonparametric methods”. Chapter of the 4. Handbook of Econometrics, North Holland, 38,2295-2339.
168- Hardle, W and Mammen, E. (1992), “Comparing Nonparametric Versus Parametric Regression Fits”.Research supported by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 123.
169- Hardle, W and Marron, J. S. (1985), “Optimal bandwidth selection in nonparametric regression function estimation”. Annals of Statistics, 13, 1465-81.
170- Hardle, W and Tuan, D. P. (1986), “Some theory on M-smoothing of time series”. Journal of Time Series analysis. 19,191-204
171- Hardle, W and Vieu, P. (1992), “Kernel regression smoothing of time series”. Journal Time series analysis, 13,209-232.
172- Hardle, W and Yang, L. (1996), “Nonparametric autoregression with Multiplicative Volatility and additive Mean”. Discussion paper, 96-62, SFB 373, Humboldt Universitât zu Berlin.
173- Hardle, W, Liitkepohl, H and Chen, R. (1996), “A review of Nonparametric Time Series Analysis”, Discussion Paper 96-48, SFB 373, Humboldt Universitât zu Berlin.
174- Hardle, W. (1989), “Applied nonparametric regression”. Cambridge University Press, Cambridge.
175- Harvey, A.C. (1988), “Forecasting Structural Time Series Models and the Kalman Filter”. Cambridge University Press.
176- Hosking, J.R.M. (1981), “Fractional Differencing”, Biometrika, 68(1), 165-176.
177- Hosking, J.R.M. (1996), “Asymptotic Distributions of the Sample Mean, Autocovariances, and Autocorrelations of Long-Memory Time Series”. Journal of Econometrics, 73(1), 261 284.
178- Hosking, J. R. M. (1981), “Fractional differencing”.. Biometrika, 68, 165–176.
179- Hosking, J. R. M. (1996), “Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series”. J. of Econometrics, 73, 261–284.
180- Hsieh, D.A. (1989), “Testing Nonlinear Dependence in Daily Foreign Exchange Rates”. Journal of Business, 62(3), 339-368.
181- Hsieh, D.A., (1991), “Chaos and nonlinear dynamics: Application to financial markets”, The Journal of Finance, Vol. XLVI, n°5, pp. 1839-1877.

182- Hurst, H.E. (1951), “Long Term Storage Capacity of Reservoirs”. Transactions of the American Society of Civil Engineers, 116, 770-799.
183- Hurvich, CM and Simon off, J.S. (1998), “Applied Smoothing parameter selection innonparametric regression using an improved Akaike information criterion”. J.R. Statist. Soc. B 60,271-293.
184- Hurvich, CM et Tsai, CL. (1989), “Regression and Time Series Model Selection in Small Samples”, Biometrika, 76, 297-307.
185- Jarque, C.M. and Bera, A.K. (1980), “Efficient test for normality homoscedasticity and serial independence of regression residuals”. Applied Statistics, Vol. 31, n°2.
186- Jarque, C.M. and Bera, A.K. (1980), “Testing the normality assumption in limited dependant. variable models”. International Economie Review, Vol. 25, n°3.
187- Jonston,.J. (1988), « Méthodes statistiques ». Paris. Economica. Tome 2.
188- Judge, G.C, Griffts, W.E, Hill, R.C, Lutkephonhl, H and Lee T.C. (1984), “The Theory and Pratice of Econometrics”. John Willy and Sons.
189- Karlsen, H.A and Tjostheim, D. (1998), “Nonparametric estimation in null recurrent time series”. Discussion paper, Quantifikation und Simulation Okonomischer Prozesse, Humboldt Universitât zu Berlin.
190- Kim, Y.T and Cox, D. (1996), “Bandwidth selection in Kernel smoothing of time series”. Journal of Time series analysis, 17, 49-63
191- Kondo, M and Taniguchi, M. (1993), “Nonparametric Approach in Time series analysis”, Journal of Time Series, 14(4), 397-408.
192- Kuan, C.M and Liu, T. (1995), “Artificial neural networks: an econometric perspective”, Journal of Applied Econometrics, Vol. 10, pp. 347-364.
193- Kugiumtzis, D, Lingiaerde, O.C and Christopher, N. (1998), “Regularized local linear prediction of chaotic time series”. Physica D, Vol. 112, pp. 344-360.
194- Kwiatkowski, D, Phillips, P, Schmidt, P and Shin, Y. (1992), “Testing the Null Hypothesis of Stationary Against the Alternative of a Unit Root: How Sure are we that Economic Time Series have a Unit Root?”. Journal of Econometrics, 54, pp. 159-178.
195- Kyrtsou, C and Terraza, M. (2000), “Is it possible to study chaotic and ARCH behavior jointly ? Application of noisy Mackey-Glass equation with heteroscedastic errors to the Paris Stock Exhange returns series”. Communication presented at the International

Conference Computing in Economics and Finance. Universitat POMPEU FABRA, Barcelona.
196- Kyrtsou, C and Terraza, M. (2000a), “Stochastic chaos or ARCH effects in stock series? A comparative study”. Proceedings of the International Conference of CEFI‘ Complex Behaviour in Economics’, in Aix-en-Provence.
197- Lardic, S et Mignon, V. (1999), « La mémoire longue en économie: une revue de la littérature ». Journal de la Société Française de Statistique, Tome 140, n°2.
198- Lecoutre, J.P et Tassi, P. (1987), « Statistique non paramétrique et Robustesse ». Economica.
199- Lewandowski, R, Newton, J, Parzen, E and Winkler, R. (1984), “The forecasting accuracy of major time series methods”. New York ; Wiley.
200- Li, W.K et McLeod, A.I. (1986), “Fractional Time Series Modelling”. Biometrika, 73(1), 217 221.
201- Ling, S. and Li, W.K. (1997). “On fractional integrated autoregressive moving-average time series models with conditional heteroskedasticity”. J. Amer. Statist. Assoc., 92, 1184–1194.
202- Ling, S.-Q., Li, W.K. (1997), “Fractional ARIMA-GARCH time series models”. J. Amer. Statist. Assoc., 92, 1184–1194.
203- Lo A.W. (1991), “Long-Term Memory in Stock Market Prices”. Econometrica, 59(5), 1279 1313.
204- Lo, A.W et McKinlay, C. (1988), “Stock Market Prices Do not Follow Random Walks Evidence from a Single Specification Test”. Review of Financial Studies, 1,41-66.
205- Lubrano, M. (1997), « Modélisation économétrique des séries temporelles non-stationnaires », GREQE-CNRS
206- Lung, G and Box, G. (1978), “On a measure of lack of fit in time series models”. Biometrika, 65,297-303
207- Mackey, M and Glass, L. (1977), “Oscillation and Chaos in physiological control systems”, Science 50, pp. 287-289.
208- Mallaris, A.G and Stein, J.L. (1999), “Methodological issues in asset pricing: random walk or chaotic dynamics”. Journal of Banking & Finance, Vol. 23, pp. 1605-1635.
209- Mandelbrot, B.B and Van Ness, J.W. (1968), “Fractional Brownian Motions, Fractional Noises and Applications”. SIAM Review. 10(4), 422-437.
210- Mandelbrot, B.B and Wallis, J. (1968), “Noah, Joseph and Operational Hydrology”. Water Resources Research, 4(5), 909-918.

211- Mandelbrot, B.B and Wallis, J. (1969a), “Computer Experiments with Fractional Gaussian Noises”. Part I, Averages and Variances, Water Resources Research, 5(1), 228-241.
212- Mandelbrot, B.B and Wallis, J. (1969e), “Robustness of the Rescaled Range RIS in the Measurement of Noncyclic Long-Run Statistical Dependence”. Water Resources Research, 5(5), 967-988.
213- Mandelbrot, B.B et Taqqu, M.S. (1979), “Robust RIS Analysis of Long-Run Serial Correlation”. Bulletin of the International Statistical Institute, 48(2), 69-104.
214- Martin, V and Sawyer, K. (1994), “Statistical techniques for modelling nonlinearities”. in Chaos and non-linear models in economic, CREEDY, J., MARTIN, V.L, Edward Edgar Publishes.
215- Masry, E and Tjostheim, D. (1995), “Nonparametric estimation and identification of nonlinear ARCH time series: strong convergence and asymptotic normality”. Econometric Theory, 11,258-289.
216- Matzner-Lober, E. (1997), « Prévision non paramétrique des processus stochastiques ». Thèse de doctorat de l'université de Montpellier II.
217- Melard, G. (1990), « Méthodes de prévision à court terme ». Bruxelles, Edition Ellipses.
218- Michel, T. (1994), « Méthodes statistiques en gestion ». Dunod, Paris.
219- Michels, P. (1992), “Asymmetry kernel functions in nonparametric regression analysis and prediction”. The Statistician, 41,439-454.
220- Mignon, V. (1997), « Marchés financiers, mémoire longue et processus chaotiques ». thèse de doctorat de L'université Paris X-Nanterre.
221- Mignon, V. (1998), « Marchés financiers et modélisation des rentabilités boursières ». Economica, Paris.
222- Mizrach, B. (1994), “Using U-statistics to detect Business Cycle Nonlinearities”. in Willi.
223- Mizrach, B. (1995), “A Simple Nonparametric Test for Independence”.
224- Nelson, D.B. (1991), “Conditional herteroskedasticity in Asset Returns: A new Approach”. Econometrica, 59, 347–370.
225- Newey, W and West, K. (1987), “A Simple Positive-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. Econometrica, 55(3), 703-708.

226- Niguez, T.-M., and A. Rubia. (2006), "Forecasting the conditional covariance matrix of a portfolio under long-run temporal dependence". Journal of Forecasting 25, 439{458.
227- Parzen, E. (1962), "On estimation of a probability density function and mode". The annals of mathematical statistics, 33,1065-1076.
228- Peligrad, M. (1987), "Properties of uniform consistency of the kernel estimates of density and of regression functions under dependence assumptions". Preprint.
229- Pham, T. D, (1986), "The mixing property of bilinear and generalized random coefficient autoregressive models". Stochastic Processes and their Applications, 23, 291-300.
230- Pham, T.D and Tran, L.T. (1985), "Some strong mixing properties of time series models", Stochastic Processes and their applications, 19, 297-303
231- Phillips, P.C.B and Perron P. (1988), "Testing for a Unit Root in Time Series Regression", Biometrika, 75, 335-346.
232- Pindyck, R. S and Rubenfled, D. (1981), "Économétrics models and economic forecasts", MC Gow HillBook Compagny, 1981.
233- Quenouille, M.H. (1949), "The joint distribution of serial correlation coefficients". Annuals of mathematical statistics, Vol. 20, 1949.
234- Racicot, F.E et Theoret, R. (2001), "Traité d'économétrie financière ». Presses de l'Université du Québec.
235- Ray, B.K. and Tsay, R.S. (1997), "Bandwidth selection for kernel regression with long-range dependence". Biometrika. (in press).
236- Rech, G, Terâsvirta, T and Tschernig, R. (1999), "A Simple Variable Selection Technique for Nonlinear Models", SFB 373 discussion paper 26,
237- Rissanen, J. (1973), "A fast algorithm for optimum linear predictors", IEEE Transactions on Automatic control, AC-18, 555
238- Robinson, M. (1977), "The estimation of a nonlinear moving average models". Stoch. Proc. And their appl. 1,81-90.
239- Robinson, P.M. (1983), "Nonparametric estimators for time series". Journal of Time series analysis, 4,185-207
240- Robinson, P.M. (1991). "Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression". J. of Econometrics, 47, 6784.
241- Roger, P. (1988), « Théorie des marchés efficients et asymétrie d'information: une revue de la littérature ». Finance, 9.
242- Rosa, M. A. C. (1993), « Prévision robuste sous une hypothèse ergodique ». Thèse de Doctorat de l'université de Toulouse I.

243- Rosenblatt, M. (1956), "A central limit theorem and a strong mixing condition". Proc. Nato. Ac. Se. USA, 42,43-45
244- Ross, S.A (1977), "Return, Risk and Arbitrage". in Friend et Bicksler (éd.), Risk and Return in finance Cambridge, Lippincott, 189-218
245- Roussas, G.G. (1990), "Asymptotic normality of the kernel estimate under dependence conditions: application to hazard rate". Journal of statistical planning and inference, 25, 81-104
246- Roussas, G.G and Tran, L.T. (1992), "Asymptotic normality of the recursive kernel regression estimate under dependence conditions". The annals of statistics, 20, 98-120
247- Roussas, G.G. (1990), « Nonparametric regression estimation under mixing conditions". stochastic processes and their applications, 36,107-116
248- Salanie, B. (1998), « Guide pratique des séries non stationnaires », Document de travail G9814, INSEE, 1998.
249- Sarda, P and Vieu, P. (1985), "Nonparametric regression estimation, application to prediction", Proceedings of the 4th E.Y.S.M., Pliska, Studia Mathematica Bulgarica.
250- Sarda, P et Vieu, P. (1988), « Vitesse de convergence d'estimateurs du noyau de la régression et de ses dérivées ». C.R.A.S., Ser. 1,306, 83-86.
251- Scheinkman, J.A and LeBaron, B. (1989), "Nonlinear Dynamics and Stock Returns", Journal of Business, 62(3), 311-337.
252- Schwarz, G. '1978), "Estimating the dimension of a model". The annals of Statistics, Vol. 6. 1978..
253- Sengupta, J.K and Zheng, Y. (1995), "Empirical tests of chaotic dynamics in market volatility", Applied Financial Economics, Vol. 5, pp. 291-300.
254- Sharpe, W.F. (1964), "Capital asset prices: A theory of market equilibrium under conditions of risk". Journal of Finance.
255- Silverman, B.W. (1986), "Density estimation for Statistics and data analysis". Chapman & Hall.
256- Sims, C.A. (2000), "Macroeconomics and reality". Econometrica, Vol. 48, 1980. SOLOW R.M., « On a family of lag distributions », Econometrica, avril 1960. THOMAS A., Économétrie des variables qualitatives, Dunod, 2000.
257- Stone, C. J. (1985), "Additive regression and other nonparametric models", Annals of Statistics, 13, 689-705.
258- Taniguchi, M and Kondo, M. (1993), "Nonparametric Approach in time series analysis". Journal of Time series analysis, 14, 397-408

259- Tjostheim, D. and Auestad, B. (1994b), “Nonparametric identification of nonlinear time series: selecting significant lags”. Journal of American Statistical Association, 89,1410-1419.
260- Thomas, R.L. (1997), “Modem econometrics”, Addison-Wesley,
261- VIVIANI, J. L. (1997), « Gestion de portefeuille », Dunod, 1997.
262- Tsai, H., and K.-S. Chan. (2007), “A note on inequality constraints in the GARCH model”, Technical Report No 361, Department of Statistics & Actuarial Science, The University of Iowa.
263- Tschernig, R and Yang, L. (1998), “Nonparametric Lag Selection for Time Series”. Journal of Time Series Analysis, forthcoming.
264- Tschernig, R. (1996), “Nonlinearities in German Unemployment Rates: A Nonparametric Analysis”, SFB 373 discussion paper 45.
265- Ullah, A. (1988), “Nonparametric estimation and hypothesis testing in econometric models”, Empec, 13,223-249.
266- Van Home, J and Parker, G. (1967), “The random walk theory: an empirical test”. Financial Analysts Journal, 23.
267- Viano, M.C, Deniau, C and Oppenheim, G. (1995), “Long-range dependence and mixing for discrete time fractional processes”. Journal of Time series analysis, 16, 323-338
268- Vieu, P. (1992), “Bandwidth selection for kernel regression”. A survey.Comp. Statist.
269- Vieu, P. (1995), “Order choice in nonlinear autoregressive models”. Statistics, OPA, 26, 307-328.
270- Watson, G.S. (1964), “Smooth regression analysis”. Sankhyâ, A26, 359-372
271- Wei, W. (1990), “Time series analysis: Univariate and multivariate methods”. Addison-Wesley
272- Weiss, A. A. (1984), “ARMA models with ARCH errors”. Journal of Time series Analysis, 5(2),129-143.
273- White, H. (1980), “A heteroskedasticity-consistent covariance estimator and a direct test for heteroskedasticity”. Econometrica, Vol. 48, 1980.
274- Willey, T. (1995), “Testing for nonlinear dependence in daily stock indices”, In Chaos and nonlinear dynamics in the financial markets, TRIPPI R.R. (ed.), Irwin, pp.105-135.
275- Wold, H. (1954), “A study in the analysis of stationary time series, Almquist-Wiksell.
276- Wonnacott, T.H et Wonnacott, R.J. (1998), « Statistique”. Economica, 4e ed., Paris, 1998.
277- Wolf, A., et al., (1985), “Determining Lyapunov exponents from a time series”. Physica 16D, 285-317.